NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

ATTRITION IN THE UNITED STATES ARMY:
AN EXPLORATORY DATA ANALYSIS APPROACH

by

David Alan Thomas

June 1984

Thesis Advisor: Peter A.W. Lewis

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Attrition in the United States Army: An Exploratory Data Analysis Approach		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis June 1984
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David A. Thomas		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943		12. REPORT DATE June 1984
		13. NUMBER OF PAGES 193
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Attrition, Enlistees, Exploratory data analysis, Graphical analysis, GRAFSTAT, Recruiting, Multivariate data, personal characteristics, Survival analysis, Survivor curves, Length of service, Education level

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Exploratory data analysis techniques were utilized to demonstrate the effectiveness of such techniques in identifying factors associated with attrition from the United States Army. Multivariate graphical data analysis was performed utilizing the "Draftsman" program recently added to the NPS GRAFSTAT package, as well as other exploratory data analysis techniques. Empirical survivor curves which take into account and explicitly display the discrete

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

probabilities of departure of enlistees at 36 and 48 months are provided. Tables are provided depicting probabilities of attrition and reenlistment for selected personal characteristics of enlistees.

Approved for public release; distribution unlimited.

Attrition in the United States Army: An Exploratory Data Analysis Approach

by

David Alan Thomas
Captain, United States Army
B.S., United States Military Academy, 1975

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 1984

ABSTRACT

Exploratory data analysis techniques were utilized to demonstrate the effectiveness of such techniques in identifying factors associated with attrition from the United States Army. Multivariate graphical data analysis was performed utilizing the "Draftsman" program recently added to the NPS GRAFSTAT package, as well as other exploratory techniques. Empirical survivor curves which take into account and explicitly display the discrete probabilities of departure of enlistees at 36 or 48 months are provided. Tables are provided depicting probabilities of attrition and reenlistment for selected personal characteristics of enlistees.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

This thesis is dedicated to the growing number of statisticians and analysts that believe in the use of Exploratory Data Analysis as a necessary part of any data analysis; in particular such people as Dr. John W. Tukey and Dr. Peter A.W. Lewis.

The graphs depicted in this thesis were produced by an experimental APL package, GRAFSTAT, which the Naval Postgraduate School is using under a test agreement with the IBM Watson Research Center, Yorktown Heights, New York. The author of this thesis is grateful to Dr. P.D. Welch and Dr. Philip Heidelberger for making GRAFSTAT available for student use.

I. INTRODUCTION

A. BACKGROUND

The inception of the All Volunteer force in 1973 provided Army manpower planners with the challenge of attracting, recruiting and retaining high quality personnel. The ever-increasing technology on the battle-field coupled with budget constraints have forced manpower planners to search for an efficient alternative to sheer numbers. The soldier of today must be able to operate and maintain highly sophisticated equipment. In addition, the Army manpower planner must also cope with a decreasing supply of 18-21 year olds. In fact, this cohort is predicted to shrink by about 15 percent by 1988 when compared to the 1979 cohort,and by about 25 per cent by 1994 [Ref. 1: p. 2].

Of course, manpower shortages in the army are nothing new. Past shortages have been both quantitative and qualitative; the shortages historically have fluctuated over the years depending on the intricate balance among military requirements, civilian employment and wage alternatives. [Ref. 1: p.1]

Currently Army recruiters have eliminated shortages. Through an extensive advertising campaign, army planners have taken maximum advantage of current economic conditions: since the inception of the All Volunteer Force, the army has met its objectives in numbers of enlistees in all but two years (FY77,FY79) and has met 100 percent of objective in the last four years [Ref. 2: pp.6-7].

The trends alluded to above, however, indicate that such ease in manning the force may be short-lived. Army manpower

planners may be forced to recruit "less-qualified" soldiers just to meet manning requirements. More screening will be necessary to meet these requirements in an adequate fashion.One particular screen that has seen widespread use is education level [Ref. 3: p.342]. Future recruiting may not result in the high percentage of High School Diploma Graduates that is currently enjoyed: from FY79 through FY83 an average of sixty percent of all non-prior-service enlistees have been High-School Diploma Graduates [Ref. 2: pp.6-7].

The FY85 Army Budget calls for holding active end strength to FY 84 levels [Ref. 4: p.16]. This in turn leads to maintaining an 80.6 per cent level of High School Diploma Graduate content [Ref. 4: p.16]. In order to maintain this level and maintain FY 84 end strength, a maximum of 12 per cent of total enlistees may be non-high-school diploma graduates (NHSDG).

In light of increased (due to inflation) or at best constant recruiting costs, army manpower planners must necessarily be concerned with determining exactly what level of education produces the best recruiting risk. In other words, if Non-High-School Diploma Graduates and Graduate Equivalent Degree enlistees are a necessary part of the force structure, what, if any, are the associations between education level and "performance"? This research effort will provide some insight to this question.

Some commonly accepted measures of performance currently in use by army manpower planners are

1. Attrition (various definitions and levels),
2. Skill Qualification Tests scores,
3. Military judicial and non-judicial actions or lack thereof.

   [Ref. 5]

The term attrition itself has taken on many different meanings in recent research. In many studies, attrition has been defined as "failure to complete the first term of service." [Ref. 6: p.24] In this study total length of service obtained will be used as a measure of performance in the initial analysis and the above definition will be used in later more detailed analysis.

Manpower policy makers have been investigating attrition since the early 60's [Ref. 7: p.1]. Such research has attempted to predict attrition through various sorts of models. Across the Army, Navy, and Air Force, level of education, mental ability and age have been determined as the best "pre-service" predictor variables of attrition [Ref. 7: p.1].

The cost of "assessing, dressing and training" a typical soldier has been estimated at approximately $15000 [Ref. 8 p.16]. This initial cost resulted in a total cost of $1,743,200,000 in reaching FY 83 enlistment goals, based on 116,215 accessions [Ref. 9: p. i]. Obviously one means of meeting requirements at minimum cost is to reduce unnecessary losses of money through premature attrition.

Attrition studies have been, for the most part, based on different forms of regression models, particularly linear and logistic models using both individual occupations and occupational groups. [Ref. 10: pp. 1-10]

B. PURPOSE OF RESEARCH EFFORT

The purpose of this thesis will be twofold:

1. To demonstrate the usefulness of Exploratory Data Analysis (EDA) techniques in "preprocessing" large volumes of data generally associated with any manpower analysis. This thesis will use a study of attrition of U.S. Army enlistees as the vehicle for

this demonstration. The dependent variable under investigation is specified as "total active federal service". This phase of the research will provide examples as to how EDA techniques can assist manpower analysts and decision makers in determining problems in the data under analysis and in variable selection. (a discussion of exploratory data analysis techniques is found at Appendix A)

2. Upon selection of suitable predictor variables of attrition, an analysis of survival functions will be utilized to provide more detailed information.

II. DATA AND METHODOLOGY

A. THE DATA

1. DMDC Cohort File Description

As stated, the data used to fulfil the purposes of this thesis was the FY79 COHORT file, maintained by the Defense Manpower Data Center (DMDC) at Monterey, California. This COHORT file is a longitudinal register of all accessions for a given year, updated at various predetermined times so as to allow for tracking of performance of that cohort in subsequent years. The FY79 cohort under investigation was last updated in September 1983. The file depicts each individual through 69 variables [Ref. 11]. FY79 was arbitrarily selected as a representative sample; it should be noted that the data from any given year may be confounded by political, social, and economic factors which are highly subjective and difficult to measure.

2. Preliminary Investigation and Data Reduction

The data set was reduced based on a request for an investigation into non-high-school-graduate performance from the United States Army Recruiting Command (USAREC), Fort Sheridan, Illinois. This request and subsequent telephonic requests for information suggested investigation into three military occupational skills (MOS): specifically at least one MOS from each of the major subdivisions of the Army, namely Combat Arms, Combat Support, and Combat Service Support. A histogram of the FY79 accessions (with non high school diploma graduate status) by MOS was developed (Appendix B) and subsequently the MOS's rank ordered by numbers accessed. Based on this ranking, Table I depicts the MOS's chosen for analysis.

TABLE I

Military Occupational Skills To Be Analyzed

Major Subgroup of Army	MOS
Combat Arms	11B Infantryman 11X Infantryman 13B Artilleryman
Combat Support	64C Motor Transport Operator 31M Multichannel Communications Operator
Combat Service Support	76Y Supply Specialist 94B Food Service Specialist

In addition, the data set was further reduced to only non-prior-service male accessions, based again on conversations with USAREC. Of course, all education levels were included so as to be able to ultimately compare the effects of education level. A data request was provided to DMDC for the data described above; the final form of the data was in character form, stored in 5 files on the Mass Storage System of the Naval Postgraduate School computer system.

3. Preparation for Exploratory Data Analysis

Based on the above reduction of the data, the 69 available variables of interest were reduced to 14 variables to limit the scope of the investigation and to demonstrate the use of the Exploratory Data Analysis techniques. It should be noted that these procedures will be useful on any size data set for any number of variables subject only to the limitations of the storage capacity of the computer system in use. Table II provides a listing of this first selection of variables.

TABLE II

Initial Variables of Interest

Explanatory Variables	Levels
High School Level Obtained	13
Current Pay Grade	31
Marital Status (current)	2
Number of Dependents (current)	9
Character of Service	4
Reenlistment Code	4
Age at Entry	15
High School Level at Entry	13
Sex	2
Race	3
Ethnic Code	20
Marital Status/No.of Dependents at Entry	20
AFQT Group (Mental Category)	8

Dependent Variable	Levels
Total Active Federal Service	Number of months

Total active Federal service was chosen as the dependent variable at this stage of the analysis to allow for investigation of possible associations with the above candidate predictor variables over time as opposed to a "go-no-go" binary representation of attrition. This dependent variable allows the decision maker to initially see the effects of the candidate predictor variables on different levels of attrition, whether the assesions contracted for three or four years of initial service.

These variables having been selected, simple FORTRAN and APL programs (Appendix C and D) were written to retrieve the data from mass storage into an interactive environment for graphical analysis.

B. METHODOLOGY

Exploratory data analysis techniques are to be utilized to analyze the data described above. A draftsman's display [Ref. 12: pp. 136,145] is prepared to initially process the data. Association between variables of interest are determined as well as any possible errors in the data. Upon necessary refinement, further Draftsman's displays are utilized to select possible explanatory variables. Boxplot analysis is performed to analyze the distribution of the levels of the candidate explanatory variables and their contribution in determining length of service. A comparison of the statistics of each distribution is utilized to determine relationships among the various levels of each of the candidate explanatory variables. Confirmatory analysis in the form of parametric and nonparametric hypothesis testing is presented to indicate the statistical significance of sample comparisons.

Finally, a survivor function approach is utilized to analyze for probabilistic relationships. Failure times and survival times are identified that lead to calculations of the probability of attrition and reenlistment for both the three year enlistees (3YO) and the four year enlistees (4YO) from the FY79 COHORT data.

III. EXPLORATORY DATA ANALYSIS

A. INITIAL DRAFTSMAN'S DISPLAY

As described in Appendix A, the draftsman's display [Ref. 12: pp. 136,145] is an efficient means of taking a "first glance" at a data set. In addition, the use of APL as a programming language for this analysis allows for rapid manipulation of large arrays in a user friendly fashion. A ten percent sample of the data set consisting of the variables listed in Table II above was run through the "draftsman" program [Ref. 13]. Sampling was performed by reading every tenth record of the data set provided by DMDC. The file was prepared in Social Security Account Number order so 10 percent sampling ensured that a country-wide sample was created. Also, the file being longitudinal, any length-biased sampling problems were avoided [Ref. 14: p.13]. The output produced was a 14 by 14 matrix of two dimensional scatter plots. A segmented copy of the display is found at Appendix E. Refer to this Appendix for the following discussion. Note that the data has been "jittered" to reduce overlap of the data or data points with the same discrete values [Ref. 12: p.21]. (note:coding of the variables is defined in [Ref. 11])

First an overall view of the entire display is very useful to an analyst in several ways:

1. Categorical data is rapidly identified by the "blocking effect" seen in most of the displayed variables, e.g., "Marital Status" vs. "MOS." This aspect of the display is critical in allowing the analyst to "see" the data and immediately determine where dummy variables, for example, may be necessary.

2. Also coding of the variables is displayed as in "Hs at entry " vs. "Total Service" One observes that the scale of education ranges from 0 to 13, with the majority of the data grouped at 4 and 13, which corresponds to 2 years of high school and a Graduate Equivalency Diploma (GED), respectively. Fortunately, in this analysis, the file description containing the coding schedule was available; one can envision the usefulness of the graphical technique in "uncoding" data sets that may not have an accompanying description. The analyst could then recode as necessary with many of the commonly-used file management systems available.
3. Errors in the data may be identified in a rapid and efficient manner. Again referring to "HS at entry" vs. "Total Service". the majority of the data is shown to be 2 year high school level (4) and GED (13). Now the official request for data to DMDC was for all education levels. Because of a simple misunderstanding and a misplaced operand in the code that extracted the data from the master cohort file, only NHSG data were provided. The use of the display, then allowed for the prevention of the costly mistake the unsuspecting analyst may have made in developing a model with erroneous data. This error-preventing aspect of this procedure manifested itself in subsequent displays as will be discussed later.
4. Although further analysis was not performed, the display allows the analyst to determine multicollinearity/interaction effects of the concomitant variables. For example, "Age at entry" vs."Number of dependents" plot may provide the impetus for further study, if either of the variables had a visual effect on, say, total service.

The use of the entire display is simple and "intuitive". It allows an analyst to bridge the gap, at least in some fashion, between the quantitative world of the analyst and the "real" world of the decision maker through the power of the brain's visual correlation abilities.

In the problem at hand, that is, the effect of various personal characteristics on performance (in the form of attrition or more generally, total service), analysis of the first column of the display is most revealing. Again, this data set has been discovered to only contain NHSG and GED; the entire spectrum of education levels will be analyzed in a subsequent draftsman's display. The first column of the display depicts scatter plots of all independent variables versus total Federal service. To aid in the discussion, seven of the plots have been reproduced in Figure 3.1 and the remaining seven in Figure 3.2 below.

Viewing both figures, no rapidly discernible or "glaring" associations are evident. This is largely a result of the sheer size of the data set.
However, much useful information is available:

1. <u>Figure 3.1, MOS Versus Total Service</u>:
   - This figure indicates that the largest number of assesions were MOS 11B (a "1" in the DMDC coding) followed by 13B(2 on coding scale). No 11X's are discernible because this entry level "basic foot soldier" MOS was not created until 1980. MOS 64C seems to have a distinct break in length of service: this break indicated that perhaps further analysis is needed. One possible explanation may be that a portion of the accessions entered this MOS only for the training, and subsequently left the service at first opportunity through an administrative discharge. MOS 76Y (6 on DMDC scale) seems to be the most successful in terms of

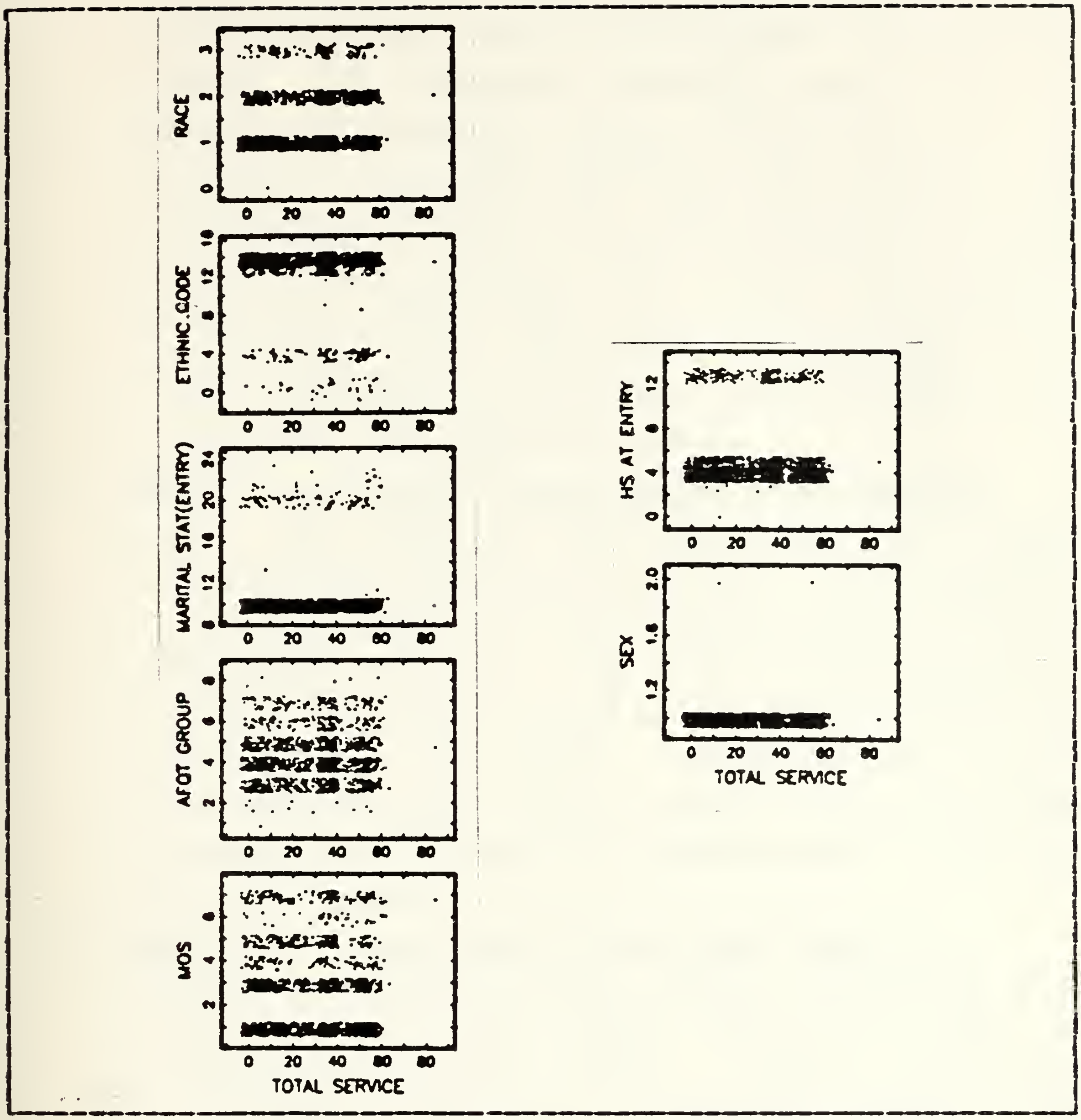


Figure 3.1 First Column of Draftsman's Display

performance (i.e., completion of at least 36 months of service).

2. <u>Figure 3.1, AFQT Group Versus Total Service</u>:
   - This variable combination is coded from 1 through 8, with 1 being a Category V. This lowest category corresponds to an AFQT total score of 9 or less. A

slight trend is evident: as AFQT increases, more data points are found in the higher total service range. This graphical analysis agrees with the literature presented in the introduction of this thesis. This also indicates that most of the accessions considered here (NHSG and GED) are found in categories 3,4,5,6, which corresponds to mental categories IVB,IVA,IIIB, and IIIA, respectively. This leads to the conclusion, as expected, that NHSDG accession performance on the AFQT is in consonance with education level.

3. Figure 3.1, Marital Status Versus Total Service:
   - This plot indicates that most NHSDG were single with no dependents(10 on the DMDC coding scheme). Again the mass of data points requires further subsequent analysis. The upper grouping depicts a slight increasing trend as accessions differ by number of dependents (20=married/no deps., 21=married/1 dep., etc.). This, along with other aspects will be analyzed in more detail in section B, this chapter.
4. Figure 3.1, Ethnic Code Versus Total Service:
   - This plot indicates that a great preponderance of the accessions were "other" which corresponds to Caucasian in the DMDC coding. Puerto Ricans (code 4) indicate increased total service, leading to the conclusion that race is an important predictor of service and attrition.
5. Figure 3.1, Race Versus Total Service:
   - This plot reinforces that of the ethnic code plot. Most assesicns were white, as indicated by the mass of data at code 1. There is a slight increase of service indicated in the category 2 corresponding to blacks, and perhaps even a greater massing in

the 50-60 month area of category 3, which corresponds to race "other". Again race seems to be a predictor of service or attrition.

6. Figure 3.1, Sex Versus Total Service:
   - Although the data requested from DMDC was for male accessions (coded 1), at least a few data points are evident in the category coded "2", corresponding to sex "female". Subsequent analysis through an APL program written to "scan" the data and count the frequency of certain data elements (Appendix D) indicated that in excess of 200 female NHSDG persons were accessed and recorded in this file. Again the draftsman's display and more generally the exploratory data analysis approach indicated erroneous data, perhaps preventing faulty analysis.
7. Figure 3.1, High School at Entry Versus Total Service:
   - Previously discussed above. This variable has seen widespread use as as a screening device in recruiting [Ref. 7: p. 1].
8. Figure 3.2, Age at Entry Versus Total Service:
   - This variable combination indicates a wide range of values because of the more "continuous" nature of the age variable. Although most NHSDG accessions were in the 18-20 year category, this plot indicates a slight increase in total service as age increases to about 25. Then the plot may indicate that older accessions do not fare as well in the measure of performance chosen. This agrees with the literature; further investigation is deemed necessary by the display.

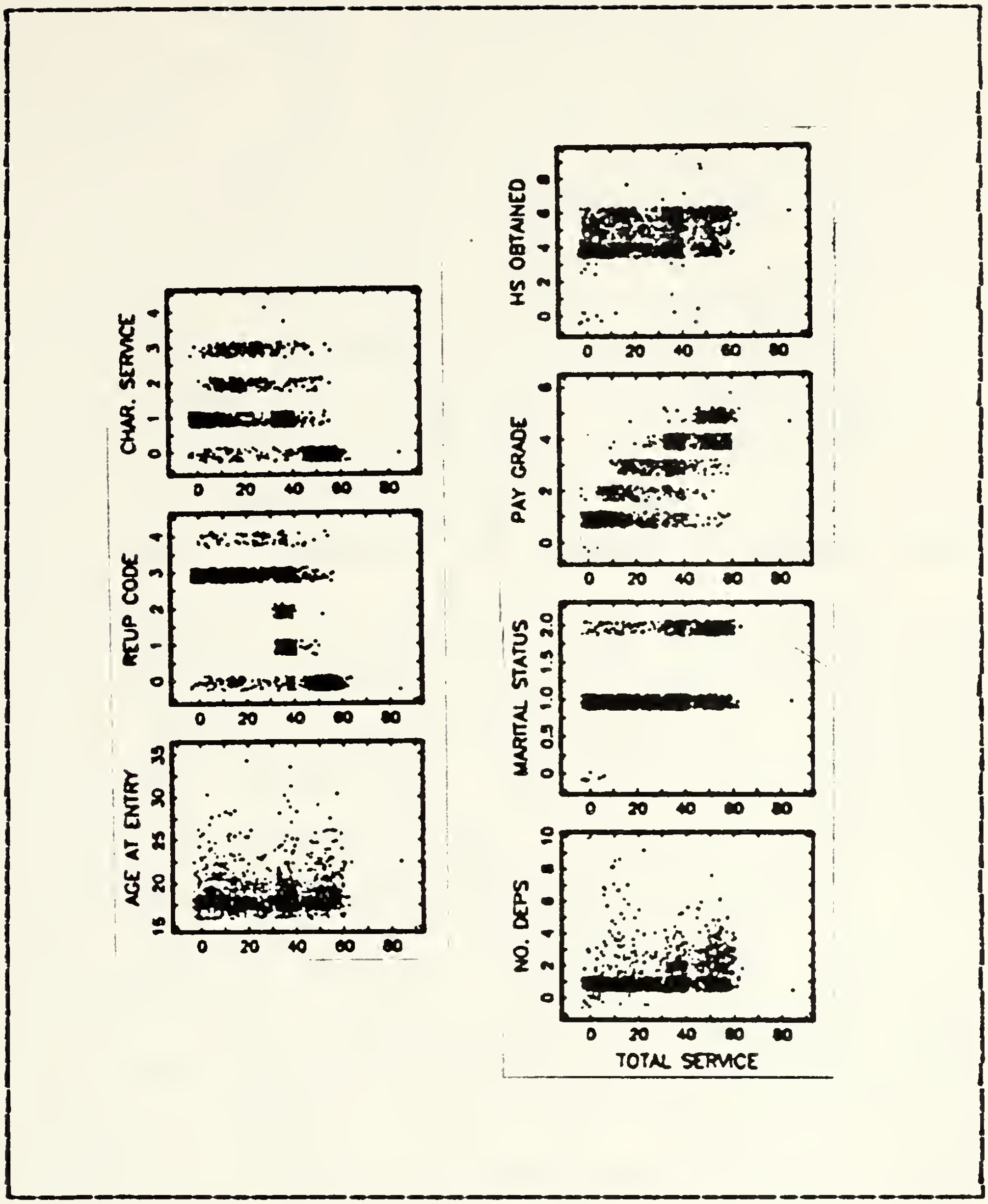


Figure 3.2   First Column Continued

9. <u>Figure 3.2, Reenlistment Code Versus Total Service</u>:
   - This plot provides information that premature attritors and successful first- tour completers may receive the same reenlistment code. A "0" indicates unknown or no code, a "1" eligibility to reenlist, a "2" that a local bar to reenlistment

has been applied, a "3" that a DA-level BAR has been applied, and a "4" that the soldier is totally ineligible. The DMDC file description only provided 2 codes: again the display has indicated erroneous data or perhaps changed codes subsequent to latest file update. This display prompted a telephonic clarification to DMDC that may not have been made in the absence of exploratory data analysis. Since both "failures" and "successes" can receive a code of "eligible", this variable is not useful as a predictor of attrition.

10. <u>Figure 3-2, Character of Service Versus Total Service</u>:
    - This variable combination, like reenlistment code above, is an indicator of "after the fact" performance, either premature attrition or successful first-term completion. Again, a "0" indicates unknown or uncoded data, a "1" an honorable discharge, "2" under honorable conditions, a "3" other than honorable conditions, and a "4" a dishonorable discharge. A massing of the data in the 0-20 month area of honorable discharges indicates that premature attritors were granted such a discharge through perhaps the Trainee Discharge Program or Expeditious Discharge Program. Thus this variable combination is not a good predictor of performance since discharge award in some cases is based on the local commander's decision. Quite often the honorable discharge may have been granted to speed the discharge of a substandard soldier; Under Honorable Conditions discharges, Other Than Honorable discharges and Dishonorable discharges require more "red tape" and administrative delays.

11. Figure 3.2, Number of Dependents Versus Total Service:

   - This plot is an updated data entry that allows for the tracking of the addition of dependents through the serviceman's tenure. The plot indicates a majority of the service people considered through September 1983 still had zero dependents (coded "1"). As number of dependents increase, total service seems to increase, although there is a proportion of data points with more dependents that failed to complete 36 months of service. This plot indicates that with further analysis, number of dependents may be a predictor of total service and attrition.

12. Figure 3.2, Marital Status Versus Total Service:

   - This variable is an updated variable indicating current marital status of the service member, a "2" for married and a "1" for other [Ref. 11]. The plot indicates a majority of the service members are still single. Married service members seem to demonstrate an increase in total service, indicating that this variable is a candidate predictor variable requiring further analysis. Again , this agrees with the literature.

13. Figure 3.2,Pay Grade Versus Total Service:

   - Although pay grade is not really a relevant means of predicting performance of accessions, this plot gives a good indication of the "intuitiveness" of the graphical depiction of the data. This plot indicates an "ideal" upward trend: as months of total service increase, the mass of data moves upward and to the right. High performers are those that receive promotions earlier than the mass in that particular pay grade, for example, grade E5

indicated by a 5 on the vertical axis of the plot shows that the majority of the service members were promoted in the 48-60 month time frame, which is the norm. However, observe the points in the 38-40 month area, indicating waiver or early promotion. From this plot, one can ascertain some idea of the number of achievers in this cohort. Along the same lines, the poor performers are evident in, for example, the E1 grade in the 50 month time frame. The entire plot seems to indicate that there is the expected upward mobility of the average soldier.

14. Figure 3.2, High School Obtained Versus Total Service:

   - Although massive data size obscures the plot, this plot does indicate the some of the NHSG accessions have completed their GED requirements (a"6" on the scale) during this time frame. Of these that have completed, a slightly increasing trend is evident, again reinforcing the literature that education level is a suitable independent or predictor variable on attrition.

In summary this first draftsman's display has demonstrated that the graphical (EDA) procedures are critical in identifying erroneous data, determining variables of interest, and identifying multicollinearity/interaction effects. Now a more refined version, with an even further reduced list of candidate variables and in some cases, with variables that have been recoded so as to be more intuitively appealing was produced. This display is in segmented form at Appendix F. This version of the display was developed from the more general data set of FY79 accessions, this time including all education levels.

B. REVISED DRAFTSMAN'S DISPLAY

The initial exploratory analysis above revealed that

1. the data on hand was not suitable in that it did not contain High School Diploma Graduates (HSDG) for use in comparing any effects of education on performance;
2. the list of variables under investigation could be further reduced.

The purpose of this second iteration of draftsman's analysis was to serve as a final check on the data and to determine any other relevant information from the data prior to a more detailed investigation utilizing other methods.

TABLE III

Reduced Variable List for Further Analysis

Dependent Variable

Total Service

Explanatory Variable

Military Occupational Skill (MOS)
Marital Status/Number of Dependents
Race
Sex
Level of Education at Entry
Age at Entry
Mental Category

General Performance Indicators

Reenlistment Code
Character of Service

The data shortfall in education level alluded to above was solved in the submission of a request to DMDC for a more complete data set. This second data set was received and

again stored in 5 files in the Mass Storage System of the NPS Computer System. The FY79 COHORT file consisted of 30778 records of data. FIles for the FY80-FY83 COHORTS were also acquired for later model validation and subsequent research. It should be noted that the target data requested was to be non-prior-service (NPS) male enlistees. The variables under investigation for this phase are listed in Table III . The last two variables in the table were, as previously stated, not to be considered as predictors but as a means of assessing general performance of the enlistees.

Again, FORTRAN and APL programs (Appendices C,D) were utilized to retrieve the data from mass storage and to manipulate it into form for interactive analysis. A ten percent sample of the data was taken for analysis (3078 records). The data was again jittered to reduce overlap.

An overall view of the draftsman's display (Appendix F) demonstrated the following:

1. The new data set is mostly categorical as expected.
2. All levels of education have been included as demonstrated by the "HS at Entry vs. Total Active Service" plot.
3. The total service scale on all plots extends to 160 months, indicating that at least some prior service enlistees have been erroneously included in the data.
4. Some female enlistees have been included as indicated by the "sex vs. Total Service" plot, again demonstrating erroneous data.

Most of the discussion below centers on the first column of the display; hence this column has been reproduced in Figure 3.3 and Figure 3.4 Viewing both figures simultaneously, the massing of the data points in heavily concentrated "blocks" demonstrates the large number of data points in the sample. The dimensionality of the problem is graphically evident. Some specific information that can be gleaned from this display follows.

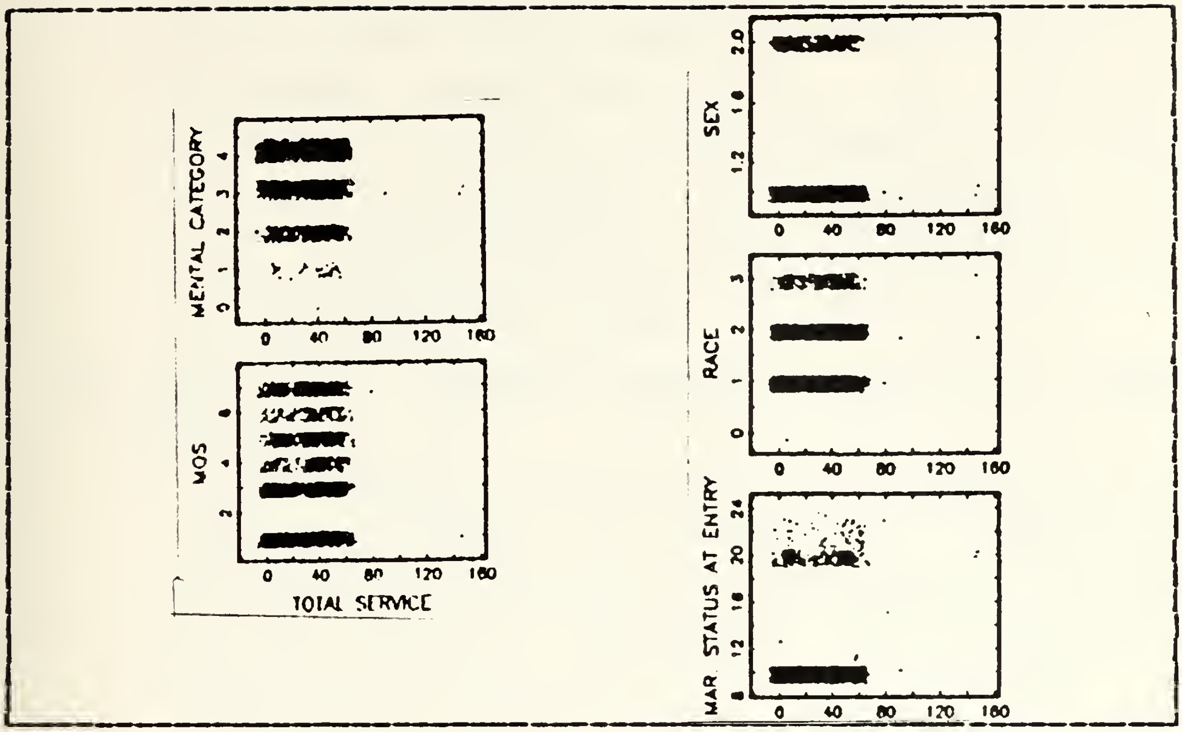


Figure 3.3 First Column of Revised Draftsman's Display

1. Figure 3.3, Military Occupational Skill vs. Total Service:
   - The 11B and 13B occupational skills (1 and 3 in the DMDC coding) appear to have attracted the most enlistees. No discernible trend or association is apparent in the length of service for these two combat arms occupational skills. Skill 64C (coded 4) demonstrates the "break" in length of service just as in the non-high-school-graduate data in the previous section. Occupational skill 76Y, (6 on the DMDC scale) as before, appeared most successful with less soldiers attriting within 36 months of service, or the normal first tour length.
2. Figure 3.3, Mental Category vs. Length of Service:
   - This variable corresponds to "AFQT group" in the original display, and has been renamed based on the

information seen in the first draftsman's display. This variable has also been recoded to reflect a 1 for mental categories 5 and 4c, a 2 for mental categories 4b and 4a, a 3 for categories 3b and 3a, and a 4 for categories 2 and 1. This plot demonstrates the recruiting policy of targeting the higher mental categories by the massing of the data in those respective areas.

3. Figure 3.3, Marital Status at Entry vs. Total Service:
   - This variable is coded with a 10 for single with no dependents through 19 for single with 8 dependents, and then a 20 for married, no dependents, through 29, for married 8 dependents. Most enlistees in this data file were single with no dependents. The married enlistees (20 and higher on the DMDC scale) indicated a slightly increased total service, particularly as number of dependents increased (21,22 and 23 on the scale)
4. Figure 3.3, Race vs. Total Service:
   - The addition of the high-school-diploma graduates has not affected the pattern that was evident in the original draftsman's display: a preponderance of whites (coded 1) followed by blacks (coded 2), and others (coded 3) is still evident. The "other" category still demonstrates increasing length of service.
5. Figure 3.3, Sex vs. Service:
   - As previously stated, this plot demonstrates that the supposed all male non-prior-service file has included in it a number of female enlistees. Also note that the massing of the data indicates that a large proportion of these females attrite with less than twenty months service.

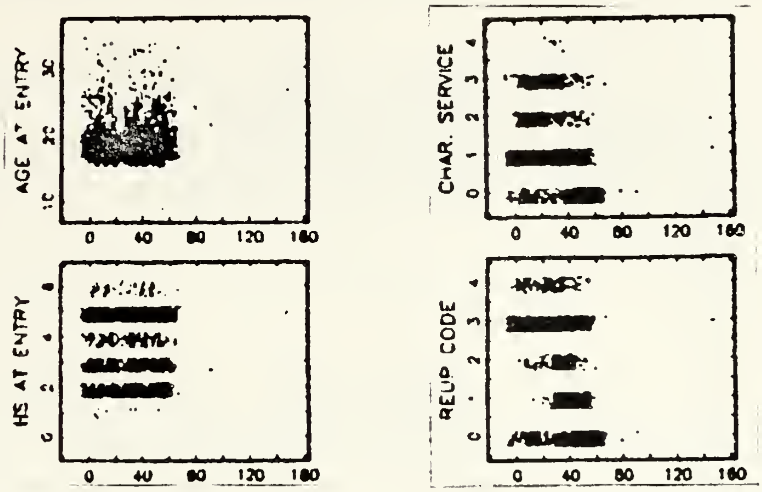


Figure 3.4 First Column, Continued

6. <u>Figure 3.4, High School at Entry vs. Total Service</u>:
   - The inclusion of all levels of education can now be verified. The plot has been recoded utilizing an APL program (See Appendix D) as shown in Table IV below: The massing of the data at position 5 indicates that most enlistees in this sample were high school diploma graduates. Also, enlistees with two to four years high school outnumber those with equivalency status. No discernible differences can be observed in the equivalency certificate holders over the other levels of education due to the massing of the data. Note the distinct break, though, in the GED length of service. A grouping is evident for zero to twenty months of service,

TABLE IV

Recoding of Education Levels for Draftsman's Display

1..... Up to 1 year of high school
2..... 2 years high school
3..... 3-4 years high school, no diploma
4..... Graduate Equivalency Diploma
5......High School Diploma Graduate
6......At least 1 year college and higher

and another for approximately thirty tc forty months. This may be partially explained by viewing age at entry versus high school at entry along the same row as high school versus total service in Figure 3.5. This same break is evident: the

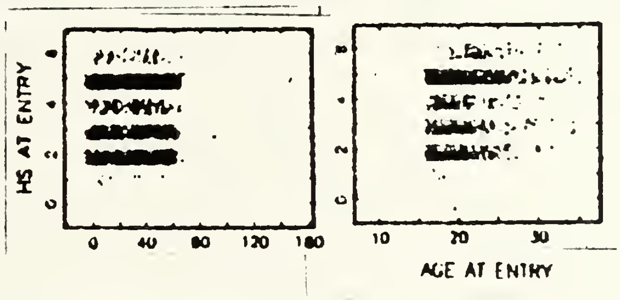


Figure 3.5 Row of Display for Cross Comparison

majority of the enlistees with the equivalency certificate (GED) are 20 years old and below with a break below 20. Perhaps this break in age when viewed with the break in service indicate that the younger GED enlistee is not as successful in completing service, just as he was in not completing high school.

7. Figure 3.4, Age at Entry vs. Total Service:

   - The massing of the data indicates that most enlistees fall in the 17-21 category. As age increases, total service again seems to be divided into distinct branches. These branches could possibly be related to marital status, number of dependents, or education level, where breaks of this nature were also evident. See the entire display to compare each of the above mentioned variable combinations versus length of service and against other possible combinations of other variables for possible insights.

8. Figure 3.4, Reenlistment Code("Reup" code) vs. Total Service:

   - This plot shows that a surprising number of enlistees, both premature "leavers" and successful (i.e., with more than 36 months service) "stayers", have an uncoded 0 reenlistment code. This probably indicates that the record has not been posted with this information. Code 1, corresponding to "eligible for reenlistment", is massed after 36 months, indicating that completion of at least the first term is a requirement for reenlistment eligibility. Code 2, a local bar to reenlistment, is also massed around 36 to 40 months of service, indicating that the decision to allow reenlistment is often reserved for the end of a soldier's term of service. It should be noted that the bar to reenlistment can be issued by the local commander at any time deemed necessary, yet it appears that exercising of this powerful option is not being done. Code 3, the Department of the Army bar to reenlistment, indicates a uniform massing of the data. This agrees with Department of the Army

policy to automatically bar soldiers that have received certain recognition as substandard through judicial, non-judicial, and administrative actions. Cross comparison with the next column of the display (see Figure 3.6) reinforces this idea: This plot is reenlistment code versus character of

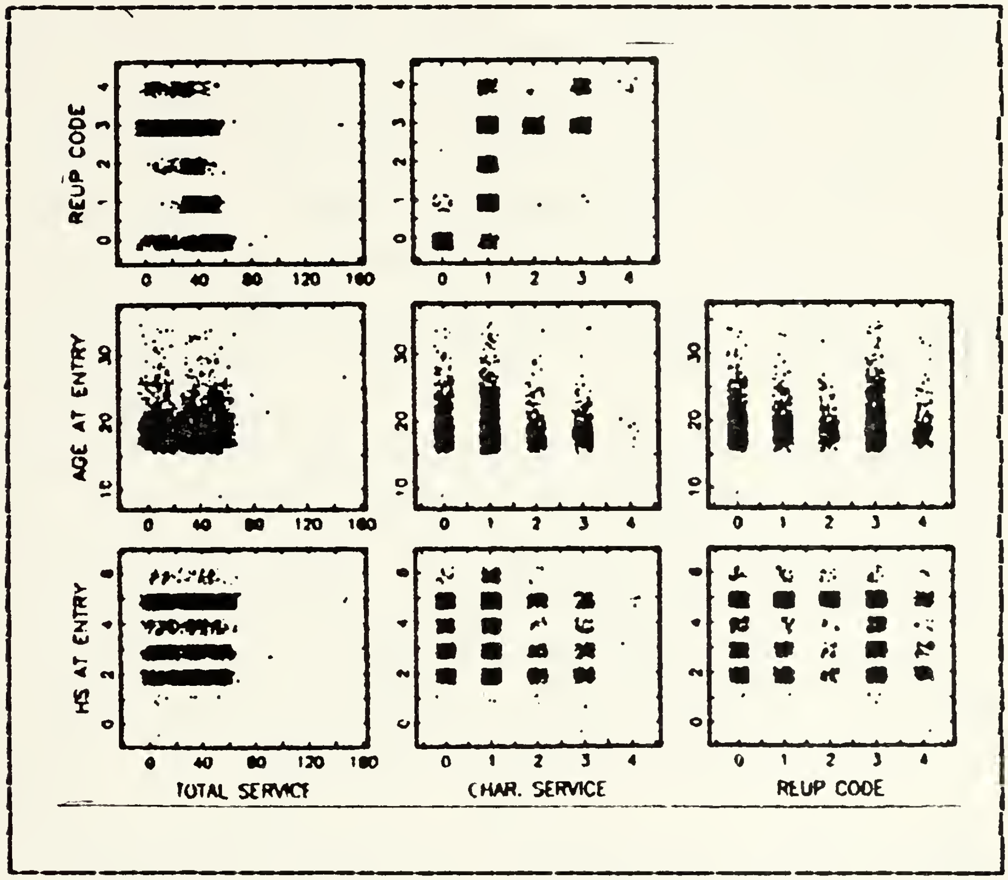


Figure 3.6 Segment for Cross-Comparison

service awarded. Reenlistment code 3 has the majority of the character of service codes 2 and 3 corresponding to "under honorable conditions" and "other than honorable " discharges respectively, both of which are considered substandard. Similarly, code 4 in reenlistment signifies

ineligibility for reenlistment, which is noticeably less of the data, contains the worst character of service code, a dishonorable discharge indicated by a 4 on the character of service axis.

9. <u>Figure 3.4, Character of Service vs. Length of Service</u>:
   - This plot indicates that as length of service increases to 38-40 months, character of service code decreases. The DMDC coding [Ref. 11] lists the coding as 1 for an honorable discharge, so one can infer that "stayers" generally receive honorable discharges. Those unsuccessful soldiers with less than 36 months service receive the majority of the "under honorable conditions", "other than honorable" and "dishonorable" discharges, coded 2,3,and 4 respectively.

This revised draftsman's display analysis has provided some insights into those personal characteristics affecting length of service and hence attrition. All seven possible explanatory variables under consideration have demonstrated at least some effects on length of service. The effects of education level, mental category, marital status and number of dependents, and age have at least initially seen to be most profound.

These displays have served to provide an initial graphical view of the data that is both intuitively pleasing and simple. In addition to identifying possible erroneous data entries and any peculiar coding of the variables, a major result of the analysis has been a reduction in the dimensionality of the data itself. The displays also aid in a general familiarization with the data under investigation.

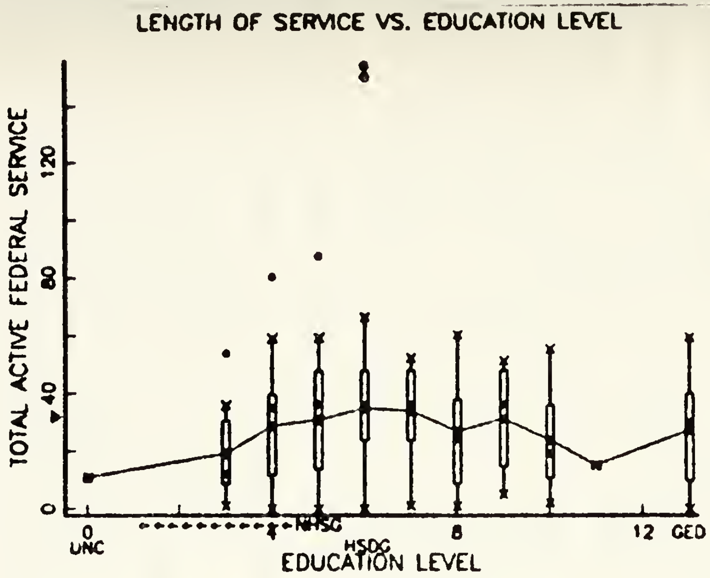
C. BOXPLOT ANALYSIS

The boxplot will be utilized to demonstrate another Exploratory Data Analysis technique that provides a powerful means of obtaining more information about a data set. (See Appendix A for a discussion of boxplots in general.)

1. Education Level Versus Length of Service

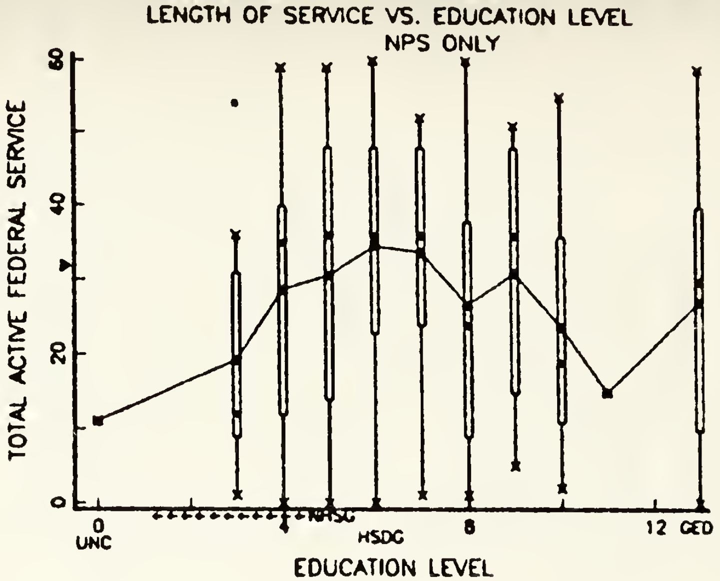
In order to demonstrate the effectiveness and power of boxplots, one of the six candidate explanatory variables, high school education level at entry, will be investigated in detail. As stated in the introduction, the literature has pointed out that education level at time of entry has been a generally accepted predictor variable in attrition analyses during the last ten years. Therefore a boxplot analysis utilizing the category subpopulation analysis program in the IBM experimental GRAFSTAT package is presented in the following discussion. Length of service, the chosen dependent variable for this portion of the study, is plotted versus education level at entry.

The left panel of Figure 3.7 is a depiction of a ten per cent sample of FY79 enlistees as of September 1983. As previously stated, this data set is a smaller subset consisting of seven military occupational skills from the entire FY79 COHORT file from DMDC. The GRAFSTAT input screen for this program allows the analyst to subdivide this batch of data into its seven component occupational skills through the input of a simple "category" vector. (See Appendix G for a depiction of this input screen.) The y axis is length of total active service in months, while the x axis indicates education level at time of entry into the army. The non high school diploma graduate level is indicated by "NHSDG" and indicates three to four years of high school without a diploma. Education level decreases toward



CATEGORY VECTOR	:EDCAT	X LABEL	:EDUC
Y	:Y←A79MOD[;1]	Y LABEL	:TOTA
SELECTION	:ALL	NO. OF ELEMENTS	:3078

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	31.875	17.638	16	36	48
0	1	0.00032489	11	0	11	11	11
3	10	0.0032489	18.2	15.883	8	12	31
4	752	0.24431	28.754	17.72	12	35	40
5	504	0.16374	30.746	17.983	14	36	48
6	1546	0.50227	34.788	17.338	24	36	48
7	22	0.0071475	33.727	15.354	24	36	48
8	20	0.0064977	26.7	18.36	8	24	38
9	11	0.0035737	31	15.13	15	36	48
10	15	0.0048733	23.733	15.823	11	16	36
11	1	0.00032489	15	0	15	15	15
13	198	0.063678	27.173	18.158	10	30	40



CATEGORY VECTOR	:EDCAT	X LABEL	:E
Y	:Y←A79MOD[;1]	Y LABEL	:T
SELECTION	:(A79MOD[;1]≤60)	NO. OF ELEMENTS	:3

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	16	36	48
0	1	0.00032541	11	0	11	11	11
3	10	0.0032541	18.2	15.883	8	12	31
4	751	0.24439	28.686	17.633	12	35	40
5	503	0.16368	30.634	17.835	14	36	48
6	1543	0.50212	34.616	16.814	23	36	48
7	22	0.0071591	33.727	15.354	24	36	48
8	20	0.0065083	26.7	18.36	8	24	38
9	11	0.0035796	31	15.13	15	36	48
10	15	0.0048812	23.733	15.823	11	19	36
11	1	0.00032541	15	0	15	15	15
13	[illegible]	0.063781	27.173	18.158	10	30	40

Figure 3.7 Education Level vs. Length of Service,I

the origin, with a 4 indicating 2 years of high school, a 3 indicating 1 year of high school, a 2 indicating 2 years of junior high school, and a 1 indicating 1 year of junior high school. To the right of the position marked "NHSDG", education level increases until the position marked "GED" which indicates a graduate equivalency "diploma" or certificate. A 6 indicates high school diploma graduate, a 7 indicates 1 year of college, 8 for 2 years of college, 9 indicates 3 to 4 years college but no degree, and a 10 indicates a college degree. Eleven and 12 indicate a masters and doctorate degree, respectively.

Viewing the entire display in the left panel, the boxplot provides a graphical statistical summary of the distribution of each of the subcategories (i.e., the levels of education) in a form for easy comparison. A table of values is also presented beneath the display. The mean of each subcategory is depicted by the dot (with lines connecting the subcategories). The median of each subcategory is depicted by the other dot in the body of the box. Adjacent values and their associated "whiskers", or the lines drawn from the body of the box to the adjacent values depict the tails of the distribution of each subcategory. Outliers are depicted by heavy dots and are defined as those values greater than 1.5 times the interquartile range of the distribution. The mean of the entire display across all subcategories is depicted by the arrowhead on the y axis.

It is immediately apparent that this subpopulation contains some enlistees with prior service, indicated by the outlier values in the left panel of figure 3.7 that have more than 60 months of service. (If an enlistee entered the Army on 30 September 1978, max length of service without prior service is 60 months through 30 September 1983.) The data requested from DMDC was to be non-prior service (NPS) only; again the important error-indication quality of this

technique is readily apparent. The non-prior service enlistees can be easily extracted from the data set in the interactive mode of this package; by the insertion of a simple truth statement (e.g., length of service ≤60 months) the data set is reduced to the desired set. (See Appendix G for the depiction of the input screen; selection is entered in the "selection" area.) This feature is also most useful in the comparison of certain other characteristics of the data as will be demonstrated.

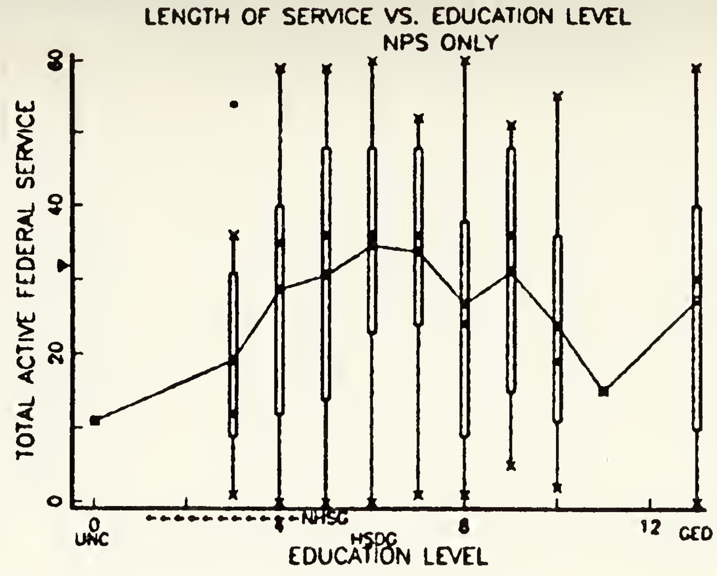
The ability to interactively subdivide the data according to other variables of interest is an extremely powerful tool, allowing the analyst to "compose" his areas of comparison and graphically determine associations in a multivariate sense. The actual data is never altered and there is no delay due to timely resubmissions of programs. In addition the graphical displays are rapidly available and intuitively appealing , requiring little explanation to those decision makers with less background in classical data analysis techniques.

The right panel of Figure 3.7 indicates the results of removing the outlier enlistees with prior service. The scale of the boxplots is now larger so the mean, median, and spread can be readily ascertained. Note that the mean length of service has been only slightly modified by the removal of the 5 outliers (determined by the "ALL" row of the two tables below the plots, 3078-3073), as expected. Also note that the shape of the boxes before and after the removal of the outliers is the same, indicative of the resistance of the boxplot. The right panel of the figure shows that performance in the form of length of service tends to increase as education level increases from the junior high level to high school diploma graduate status. College level enlistees and higher do not demonstrate this trend. Note that enlistees with the graduate equivalency

certificate demonstrate poorer performance than non-high-school-diploma graduates (NHSDG). The variance of the distribution of the enlistees with the GED is greater, perhaps due to the manner in which the GED may be awarded: Persons with any level of education can test for and gain the GED, and numerous tests are available for the certificate. Hence the large variance may be explained by the "variance" of the means of awarding the certificate.

In the next figure, (Figure 3.8), the non-prior-service enlistees are again presented in the left panel as a "comparative other." In the right panel of the figure, all those soldiers who had an uncoded education level and those with PhD degrees have been deleted, again utilizing the selection feature of the GRAFSTAT program. This results in a "smoother" line through the means, allowing a more discernible view of the association of each level of education and length of service. Note from the table that only 2 persons were in these deleted categories, hence they are referred to as "outlier" education levels in the title of the plot. These points will remain deleted throughout the remainder of this particular analysis.

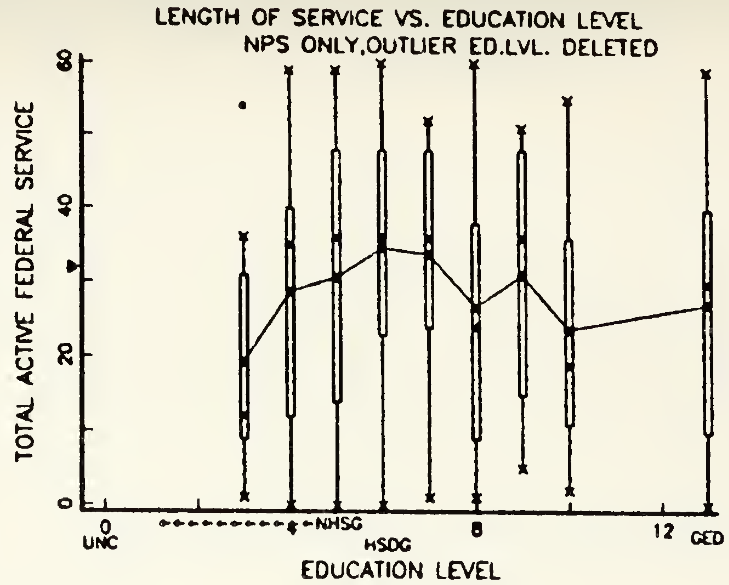
In Figure 3.9, left panel, the previously discussed boxplot with "outlier" education levels deleted, non-prior-service enlistees is again presented for further comparison. In the right panel, utilizing the selection capability, all those enlistees with any years of college have been deleted. Now the increasing trend is clearly evident. As education increases, length of service increases. Again those soldiers with the graduate equivalency certificate (GED) demonstrate a lower mean performance than both the non-high-school-graduates and the high school diploma graduates. To isolate this trend even further, GED soldiers are deleted in the right panel of Figure 3.10 Now, compared to the left panel of this display, the upward trend is most evident.



CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;1]
SELECTION :(A79MOD[;1]≤60)

X LABEL :E
Y LABEL :T
NO. OF ELEMENTS :3

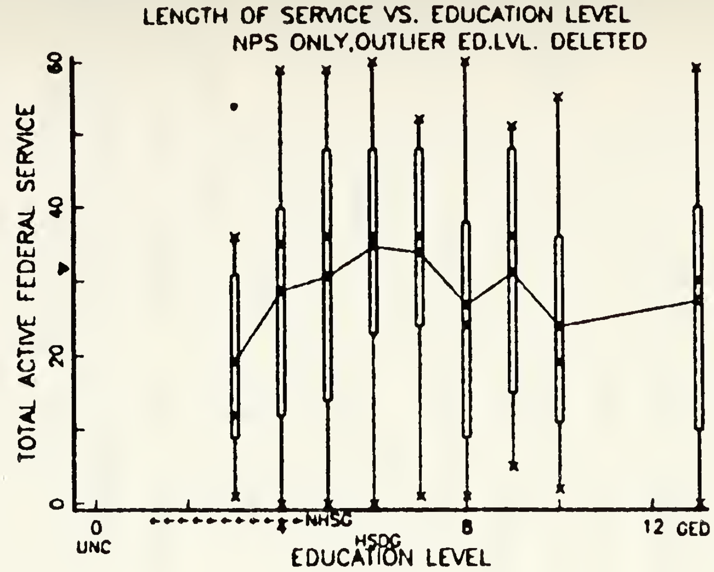
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	36	48
0	1	0.00032541	11	0	11	11	11
3	10	0.0032541	19.2	15.683	9	12	31
4	751	0.24439	28.666	17.833	12	35	40
5	503	0.16368	30.634	17.835	14	36	48
6	1543	0.50212	34.610	18.814	23	36	48
7	22	0.0071591	33.727	15.354	24	36	48
8	20	0.0065083	28.7	19.36	9	24	38
9	11	0.0035796	31	15.13	15	36	48
10	15	0.0048812	23.733	15.923	11	19	38
11	1	0.00032541	15	0	15	15	15
13	196	0.063781	27.173	18.159	10	30	40

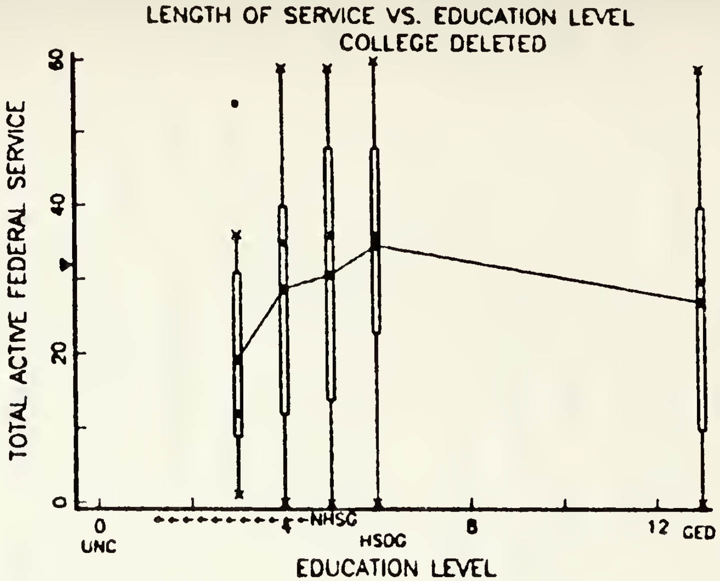


CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;1]
SELECTION :(A79MOD[;1]≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;2]

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3071	1	31.864	17.527	18	36	48
3	10	0.0032563	19.2	15.683	9	12	31
4	751	0.24455	28.666	17.633	12	35	40
5	503	0.16378	30.634	17.835	14	36	48
6	1543	0.50244	34.618	18.814	23	36	48
7	22	0.0071638	33.727	15.354	24	36	48
8	20	0.0065125	28.7	19.36	9	24	38
9	11	0.0035819	31	15.13	15	36	48
10	15	0.0048844	23.733	15.923	11	19	38
13	196	0.063823	27.173	18.159	10	30	40

Figure 3.8 Education Level vs. Length of Service,II





CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;1]
SELECTION :(A79MOD[;1]≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;2]

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3071	1	31.884	17.527	16	36	48
3	10	0.0032563	19.2	15.683	8	12	31
4	751	0.24455	28.686	17.833	12	35	40
5	503	0.16379	30.634	17.835	14	36	48
6	1543	0.50244	34.818	16.814	23	36	48
7	22	0.0071638	33.727	15.354	24	36	48
8	20	0.0065125	28.7	19.36	9	24	38
9	11	0.0035819	31	15.13	15	36	48
10	15	0.0048844	23.733	15.923	11	16	38
13	196	0.063823	27.173	18.159	10	30	40

CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;1]
SELECTION :(A79MOD[;1]≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;2

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3003	1	31.929	17.529	16	36	48
3	10	0.00333	19.2	15.683	8	12	31
4	751	0.25008	28.686	17.833	12	35	40
5	503	0.1675	30.634	17.835	14	36	48
6	1543	0.51382	34.818	16.814	23	36	48
13	196	0.065268	27.173	18.159	10	30	40

Figure 3.9 Education Level, College Deleted

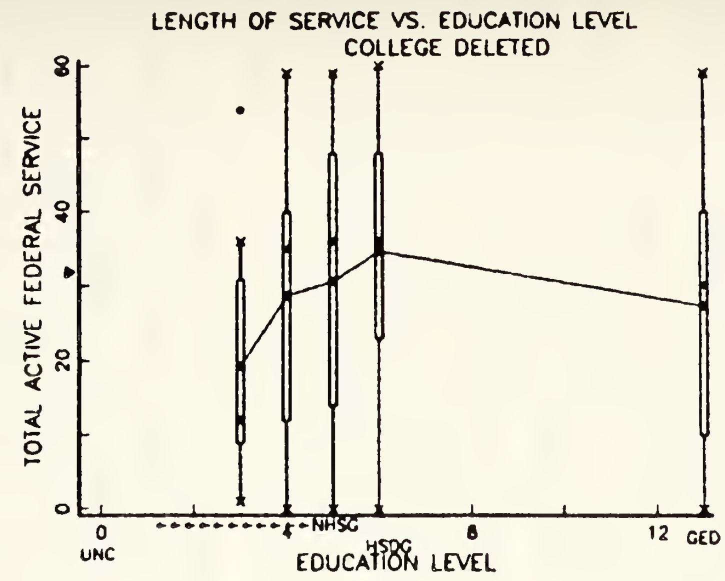
The distribution of each education level is graphically described in the boxplot. For example, in Figure 3.10, the high-school-diploma-graduate distribution appears to be symmetric as indicated by the position of the mean and median relative to the ends of the body of the box. The distribution appears to have a "thicker" tail in the lower length of service side, depicted by the longer whisker on the bottom of the boxplot. Rapid comparisons of the distributions of each of the subcategories can be done through the boxplot analysis.

Thus the use of boxplots has indicated that level of education does have an effect on performance in the form of length of service. It should be reiterated that this phase of the analysis is exploratory in nature; the comparison of means could be strengthened through such confirmatory analysis as a one-way ANOVA or in a non-parametric test such as the Kruskal-Wallis test of equality of means. An example of this confirmatory analysis will be provided in a later section of this thesis.

2. <u>The Multivariate Capability of the Boxplot</u>

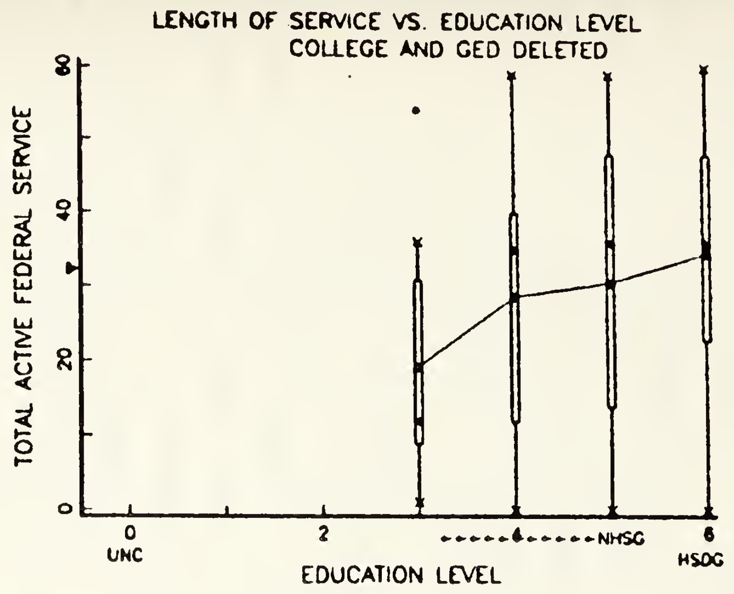
Military occupational skill of the enlistee will be examined for any indication of an association with performance, and then further analyzed with regards to education level within the occupational skill. Seven military occupational skills (MOS) were considered in this analysis. (Refer to Table I, previous chapter). Again a ten percent sample of the data was utilized for analysis.

Figure 3.11, left panel, provides the first look at this subcategory of the data. Military occupational skill is plotted against total active service. Each military occupational skill is presented on the x-axis, beginning with combat arms (11B, 11X, 13B), combat support (31M, 64C),



CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;1]
SELECTION :(A79MOD[;1]≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;2

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3003	1	31.029	17.520	16	36	48
3	10	0.00333	19.2	15.883	8	12	31
4	751	0.25008	28.666	17.833	12	35	40
5	503	0.1675	30.834	17.835	14	36	48
6	1543	0.51382	34.816	16.814	23	36	48
13	196	0.065268	27.173	18.159	10	30	40



CATEGORY VECTOR .EDCAT
Y :Y←A79MOD[;1]
SELECTION :(A79MOD[;1]≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;2]

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	2807	1	32 261	17.436	17	36	48
3	10	0 0035625	19 2	15.883	8	12	31
4	751	0.26755	28 666	17.833	12	35	40
5	503	0.17919	30.634	17.835	14	36	48
6	1543	0.5497	34.816	16.814	23	36	48

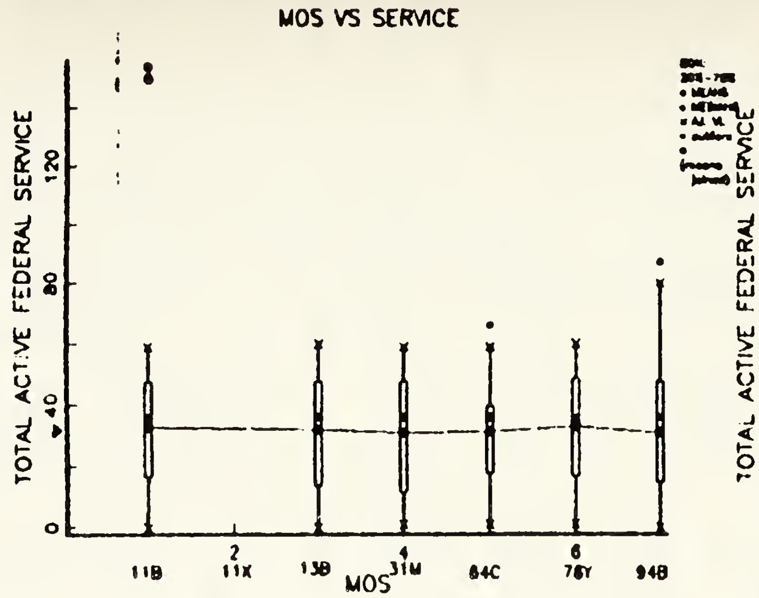
Figure 3.10 Education Level, GED Deleted

and combat service support (76Y,94B). As before, the distribution of each military occupational skill with respect to total service is presented by the boxplot. Again note the outliers in total service; these same 5 soldiers have been deleted in the right panel so as to expand the scale of the boxplots for analysis. Through the graphical depiction of these distributions, the occupational skill most successful in terms of service can be readily determined. Here, skill 76Y, supply specialist, demonstrates the greatest success followed by 11B, Infantryman.

The spread of the distribution can be observed in the body of the box; note that the deviation is also provided in the table below. The 11X occupational skill is noticeably vacant. This MOS was created as an entry level of infantryman in FY80, hence this data set contains none of this general skill. Comparing the three branches of skills described above, combat arms was most successful in performance as determined by viewing 11B and 13B boxplots collectively. Combat service support was next in order of months successful service followed by combat support.

The shape of each of the distributions is readily apparent in the boxplots. Each military occupational skill appears to be skewed toward the lower range of service as indicated by the postion of the median inside the body of each boxplot. The variances appear relatively constant except for skill 64C. The utility of the boxplot in determining homogeneity of variance, for example, for subsequent regression analysis or for ANOVA assumptions is readily obvious.

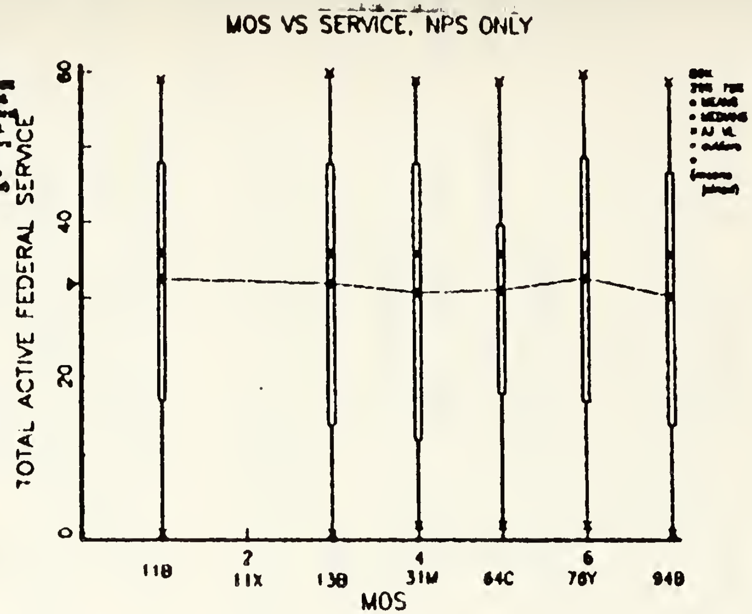
In Figure 3.12, left panel, the previous figure of non-prior-service-only enlistees has been reproduced to allow comparisons. In order to isolate the effects of education level, all high-school-diploma graduates (and 60 enlistees with at least some college) have been selected out of the subpopulation, using the selection capability. Thus



CATEGORY VECTOR :CATMOS
Y :Y←A78MOD[;1]
SELECTION :ALL

X LABEL :MO
Y LABEL :10
NO. OF ELEMENTS :30

CATEG.	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	50	.75
ALL	3078	1	31 975	17 838	16	36	48
1	1080	0.35088	32 78	18 514	17	36	48
3	702	0.22807	32	17 821	14	36	48
4	233	0.075699	30 906	18 048	12	36	48
5	433	0 14068	31.376	16 156	18	36	40
6	185	0 060104	32.881	17 513	17	36	49
7	445	0 14457	30.786	17.656	15	36	48



CATEGORY VECTOR :CATMOS
Y :Y←A78MOD[;1]
SELECTION :Y<60

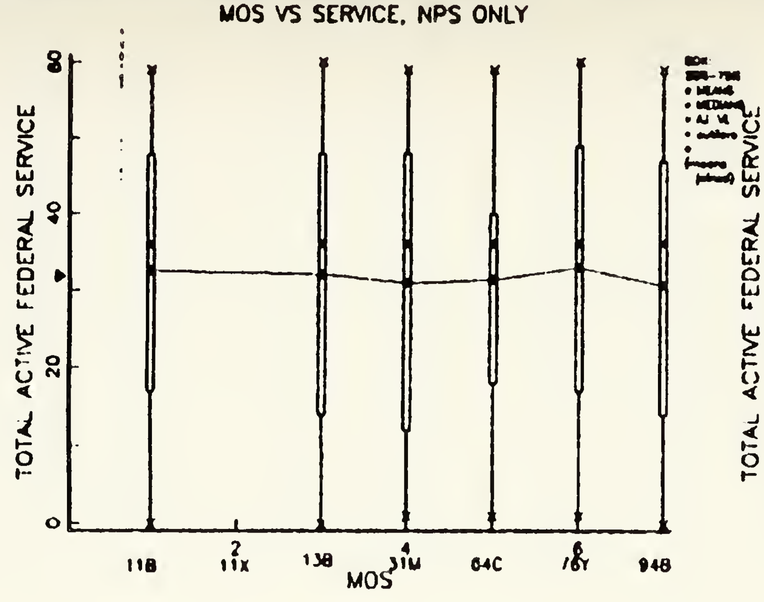
X LABEL :MO
Y LABEL :10
NO. OF ELEMENTS :30

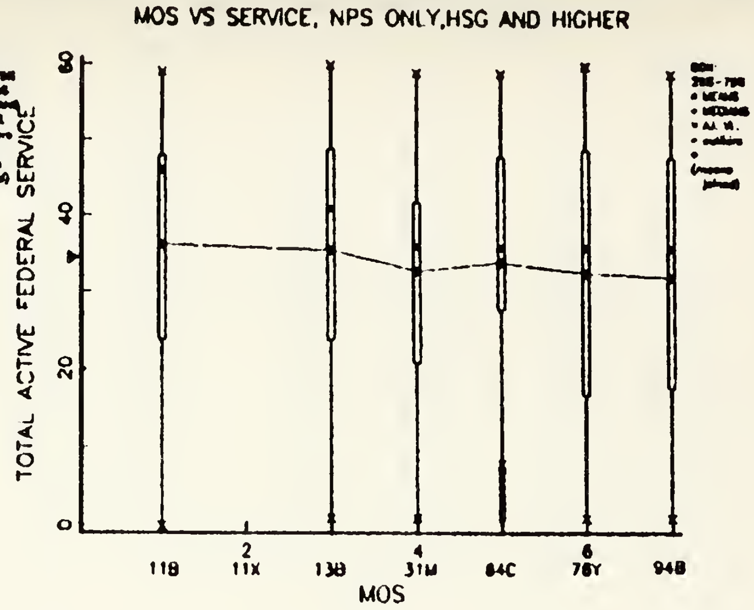
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	16	36	48
1	1078	0 3508	32 558	17.804	17	36	48
3	702	0 22844	32	17 821	14	36	48
4	233	0 075822	30 906	18.048	12	36	48
5	432	0 14058	31.745	16.088	18	36	40
6	185	0.060202	32 881	17 513	17	36	49
7	443	0.14416	30 558	17 335	14	36	47

Figure 3.11 Military Occupational Skill vs. Length of Service

in the right panel, non-prior-service, high school diploma graduates have been plotted against total active service. Immediately note that a higher mean service is observed by these better educated enlistees. The variance of occupational skills has been reduced indicating that this subpopulation is not as great a risk when considering how long each will serve. Those soldiers in skill 11B demonstrate the highest success (mean 36.15 months). Based on these graphical results high school diploma graduates perform better than the average non-prior-service enlistee, and more specifically, combat arms skills rank first followed by combat support, and then combat service support. This may seem counterintuitive as technical skills required are thought to be higher in the combat service support branches.

Figure 3.13 repeats the same sort of analysis, this time comparing non-prior service enlistees total service in the left panel to non- prior-service enlistees that have not received a diploma (NHSG) and those that have received a graduate equivalency diploma (GED). Note that "success" has fallen from 31.85 months to 29.08 months. Occupational skill 76Y has the highest mean success, and combat service support now leads in total mean service. Also note that occupational skill 31M demonstrates a much higher variance when only these non-high- school graduates are compared to the total population. This could perhaps be explained by the lack of any real standards in how the GED is awarded. This certificate can be awarded at any level of education, provided that one of numerous tests has been passed. Also the non-high-school graduate with ,say, only a 10th grade level of education could account for extremely poor performance in this somewhat technical microwave operator skill. On the other hand, those enlistees with an equivalency degree or 11-plus years of education from a "technically





CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION :Y≤60

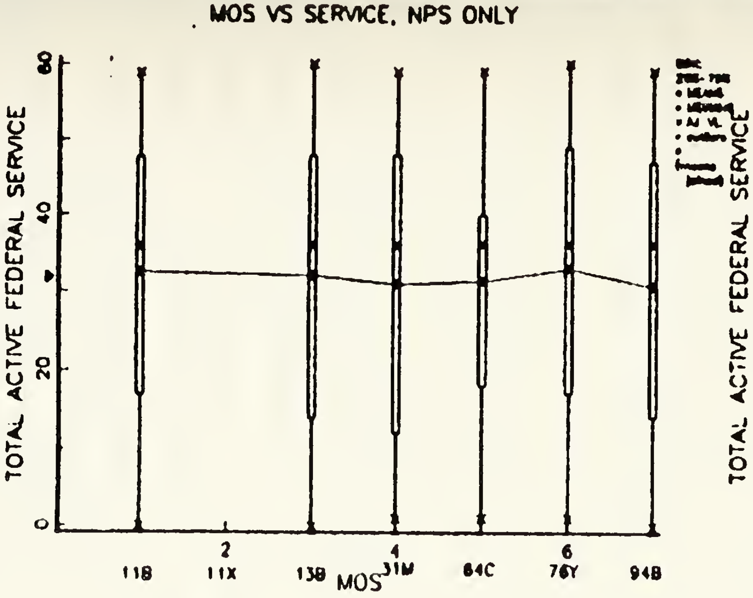
X LABEL :MO
Y LABEL :TO
NO OF ELEMENTS :30

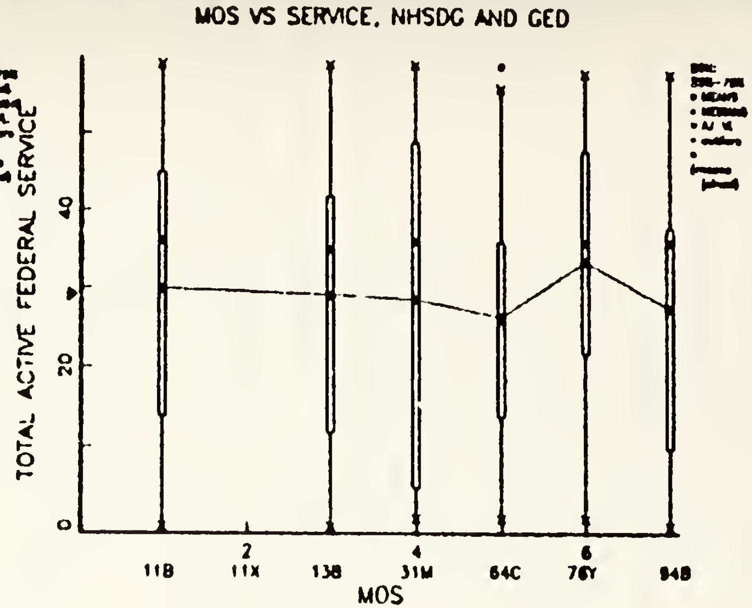
CATEG.	NO.PTS	o/o PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31 852	17 528	16	36	48
1	1078	0.3508	32 558	17 804	17	36	48
3	702	0.22844	32	17 821	14	36	48
4	233	0 075822	30 906	18 048	12	36	48
5	432	0 14058	31 245	16 088	18	36	40
6	185	0 060202	32 881	17 513	17	36	49
7	443	0.14416	30 558	17.335	14	36	47

CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;2]≥6)∧(B79MOD[;2]<13)

CATEG.	NO.PTS	o/o PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34 387	18 869	22	36	48
1	463	0.28722	36 149	17 189	24	46	48
3	315	0.19541	35 524	17.705	24	41	48
4	128	0 079404	37 805	15 755	21	36	42
5	277	0 17184	33 857	15.388	28	36	48
6	143	0 08871	32 671	17.482	17	36	49
7	286	0.17742	32 154	18 527	18	36	48

Figure 3.12 Occupational Skill,High School Graduates





CATEGORY VECTOR .CATMOS
Y :Y←A79MOD[;1]
SELECTION .Y≤60

X LABEL :MO
Y LABEL :10
NO OF ELEMENTS :30

CATEG.	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	36	48
1	1078	0.3508	32.558	17.804	17	38	48
3	702	0.22844	32	17.821	14	36	48
4	233	0.075822	30.906	18.048	12	36	48
5	432	0.14058	31.245	18.088	18	36	40
6	185	0.060202	32.881	17.513	17	38	48
7	443	0.14416	30.558	17.335	14	36	47

CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;2]<3)∨(B79MOD[;2]≥13)

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1461	1	29.077	17.82	12	36	42
1	615	0.42094	29.855	17.781	14	36	45
3	387	0.26489	29.132	17.396	12	35	42
4	105	0.071869	28.50	20.26	5	38	48
5	155	0.10609	26.4	18.171	14	28	38
6	42	0.028747	33.595	17.868	22	38	48
7	157	0.10746	27.65	18.364	10	36	38

Figure 3.13 Occupational Skill, NHSG and GED

oriented" school system could account for the high performance. More research could be performed in the area of "levels" of equivalency status and in a demographic analysis of those non high school graduates.

A comparison of high school diploma graduates to non-high-school-diploma graduates is presented in Figure 3.14 . Those soldiers with equivalency certificates (GED) have been included in the non high school graduate category. In the left panel of this figure, the high school graduate category is presented for comparison. Again, this plot is non prior service only graduates. The right panel is non-prior-service,non-high-school-graduates and equivalency certificate holders. The "better educated" enlistees ,on the average, have outperformed those with less than a school degree(as shown by the mean values both in the boxplots and the tabled values below the plot). Skill 76Y, supply specialist,in the right panel, indicates a better performance, (mean of 33.6) than the same occupational skill with a high school diploma (mean 32.7). The variance of the two distributions is roughly the same. Again personal characteristics and the technicality of the skill involved along with the effects of mental group distributions, enlistment bonus differences may explain this anomaly.

Since a difference has been established in performance based on the certificate status at entry, the boxplot can be utilized to go even into further detail. In Figure 3.15,non-high-school-graduates only are compared to high-school-diploma gradtates and in Figure 3.16 those with equivalency degrees are compared to the diploma-carrying soldiers. In each figure, the batch remains limited to non-prior-service enlistees, varying only education levels.

Figure 3.15 demonstrates that the non-high-school-graduate performance ,depicted in the right panel, was below diploma -graduate performance. This is seen

graphically by comparing the relative position of the mean lines and in the tabled value below each plot. Going further, a comparison of each individual occupational skill in the left panel to its counterpart in the right panel indicates that the varied educational level produces and entirely different distribution. This is observed through the location of the respective means and medians, the size of the body of the box, and the length of the"whiskers" or tails of the distribution.

In Figure 3.16, this same difference in the distribution of military occupational skills with respect to education level is again obvious. In every case, the high school diploma graduates outperform the soldiers with equivalency status. Again, GED holders exhibit a larger variance as indicated by the body of the box, indicating a higher risk in attrition.

The entire analysis presented on the effects of education level within military occupational skill is summarized in Figure 3.17 , where the baseline of non-prior service enlistees, categorized by occupational skill versus length of service is displayed simultaneously. Education level has been selected in each plot. Education level, military occupational skill, and length of service have been integrated into a single display. Any other combination of variables such as marital status, age, race, could be further selected to provide more of a multivariate display. The EDA techniques combined with the IBM GRAFSTAT package allows for any combination of covariates in an analysis, limited only by the imagination of the analyst. This display allows a rapid comparison of the effects of education level on performance (in the form of length of service), perhaps providing a strong argument in favor of these graphical methods for at least initial decisions regarding what level of education to recruit. Again,

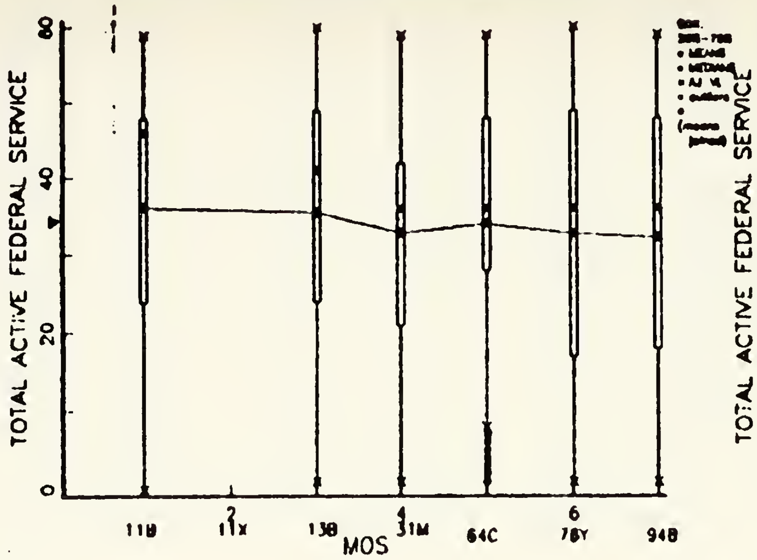
confirmatory analysis is necessary for a more refined analysis of relative merits of alternative educational policies.

3. Summary of Boxplot Analysis of Remaining Variables

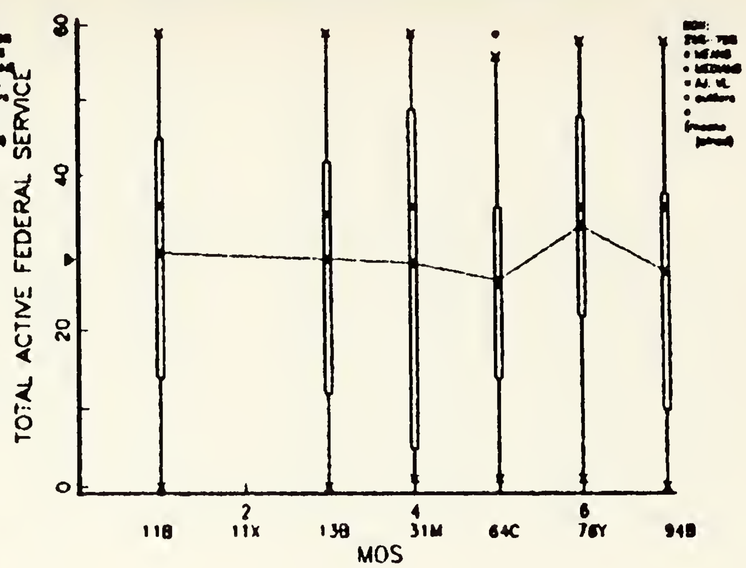
The remaining possible explanatory variables and the two "general performance indicators" were analyzed in the same manner as presented above. In each case the effects of education level were observed in a multivariate sense by the use of selection combined with the pairwise comparison of length of service and the other candidate explanatory variables. The actual boxplots are found in Appendix H. Again, confirmatory analysis should be performed to verify the statistical significance of any conclusions drawn from this exploratory analysis. A summary of this analysis is presented below:

1. Mental Category Versus Length of Service: (See Figures H.1 through H.9, Appendix H)
   - As mental category increases, length of service increases. Category 4c outperform Categories 4b,4a,3b.
   - High school diploma graduates outperform the non-high-school-graduates and equivalency certificate holders (GED) in all categories.
   - No category 1 soldiers were observed in the non-high-school-graduate or GED categories.
   - Non-high-school-graduates outperformed the equivalency certificate (GED) holders in every category.
   - Variance in the non-high-school graduate and GED performance is higher than that of the diploma graduates, indicating higher risk in attrition. Equivalency certificate (GED) holders' variance was observed to be higher than non-high-school-graduates, again indicating a higher risk.
   -

MOS VS SERVICE, NPS ONLY,HSG AND HIGHER



MOS VS SERVICE, NPS ONLY,NHSG AND GED



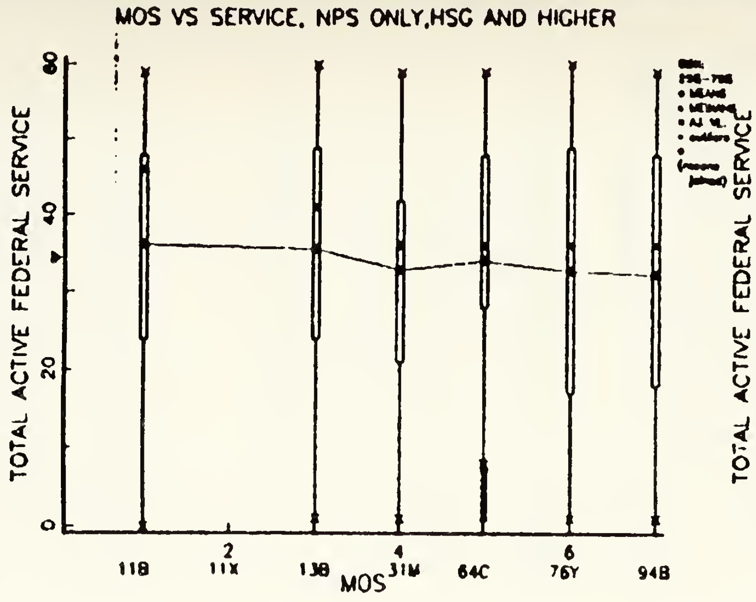
CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;2]≥6)∧(B79MOD[;2]<13)

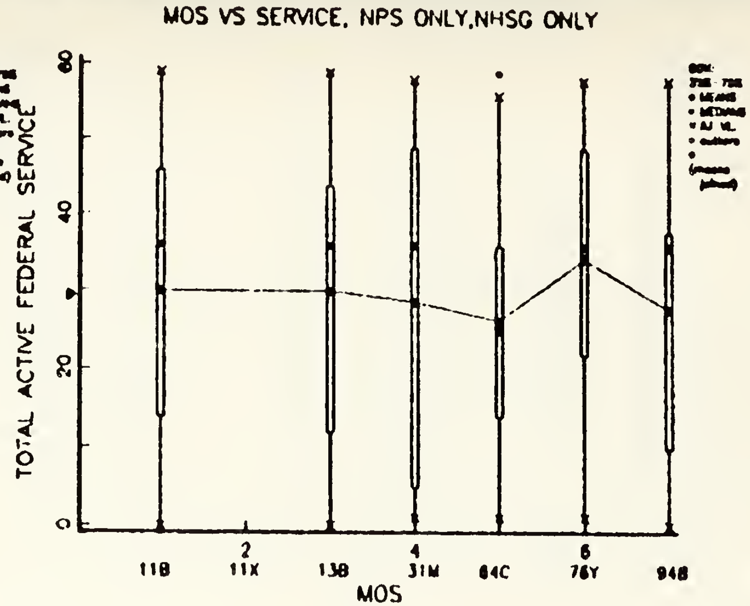
CATEG	NO PTS	0/0-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34 367	18.869	22	36	48
1	463	0 28722	36 149	17.189	24	46	48
3	315	0 19541	35 524	17 705	24	41	48
4	128	0 079404	32 805	15 755	21	36	42
5	277	0 17184	33 957	15 389	28	36	48
6	143	0 08871	32 671	17 482	17	36	49
7	286	0.17742	32.154	16 527	18	36	48

CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧((B79MOD[;2]≤5)∨(B79MOD[;2]≥13))

CATEG	NO PTS	0/0-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1441	1	29 077	17.82	12	36	42
1	615	0 42094	29 855	17.781	14	36	45
3	387	0 26489	29 132	17 398	12	35	42
4	105	0.071869	28 59	20 28	5	36	49
5	155	0 10609	26.4	18 171	14	26	36
6	42	0.028747	33 595	17 668	22	36	48
7	157	0.10746	27 65	18.364	10	36	38

Figure 3.14 HSDG, NHSG and GED by Occupational Skill





CATEGORY VECTOR CATMOS
Y :Y←A79MOD[;1]
SELECTION .(Y≤60)^(B79MOD[;2]≥6)^(B79MOD[.2]<13)

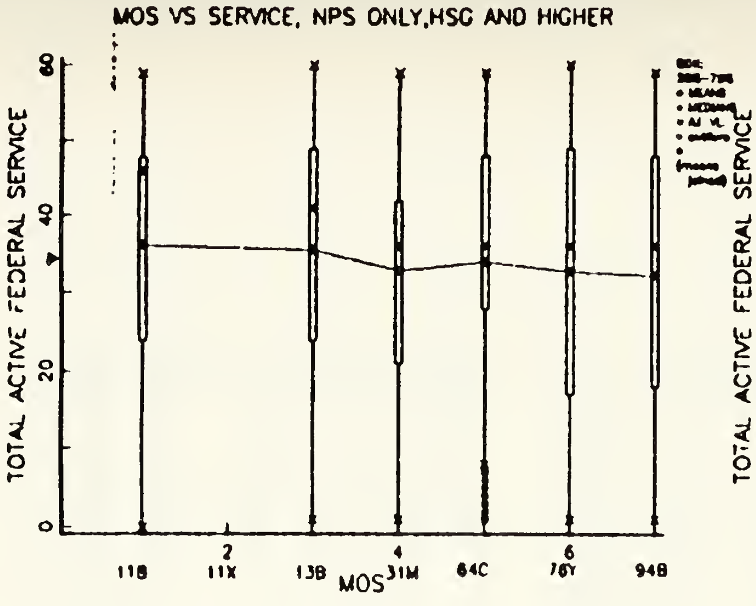
CATEG.	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34 367	18.869	22	36	48
1	463	0 28722	36 149	17 189	24	48	48
3	315	0 19541	35 524	17 705	24	41	48
4	128	0.079404	32 805	15 755	21	36	42
5	277	0 17184	33 957	15 389	28	36	48
6	143	0 08871	32 671	17 462	17	36	49
7	286	0 17742	32 154	16 527	18	36	48

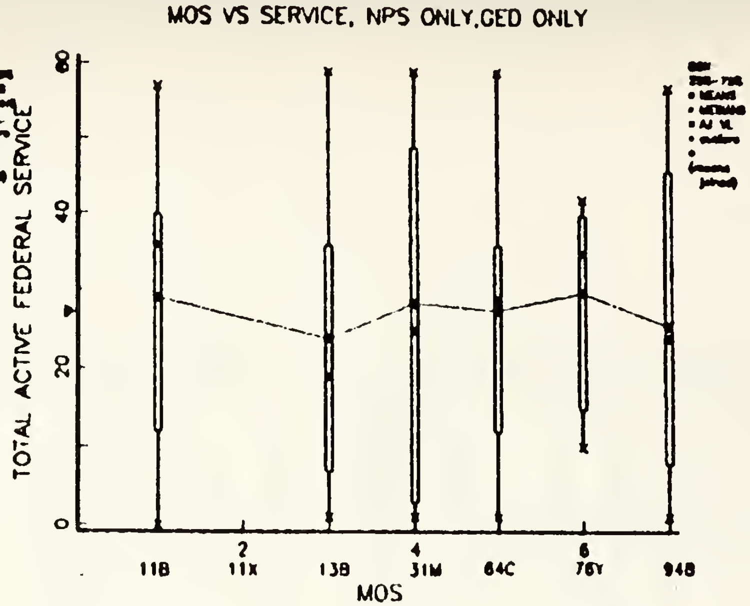
CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)^(B79MOD[.2]<5)

X LABEL
Y LABEL
NO. OF ELE

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1265	1	29 372	17.748	12	36	43
1	532	0 42055	28 872	17 795	14	36	46
3	331	0 26166	30 074	17 144	12	36	44
4	88	0 069565	28 625	19 808	8	36	49
5	136	0 10751	26 243	16 001	14	25	36
6	35	0 027668	34 343	18 504	22	38	49
7	143	0 11304	27 846	18 282	10	36	38

Figure 3.15 Non High School Grads vs. Diploma Grads





CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;2]≥6)∧(B79MOD[;2]<13)

CATEG.	NO.PTS	%-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1812	1	34.387	16.869	22	36	48
1	463	0.28722	36.149	17.189	24	46	48
3	315	0.19541	35.524	17.705	24	41	49
4	128	0.079404	32.805	15.755	21	36	42
5	277	0.17184	33.957	15.389	28	36	48
6	143	0.08871	32.871	17.462	17	38	49
7	286	0.17742	32.154	16.527	18	36	48

CATEGORY VECTOR :CATMOS
Y :Y←A79MOD[;1]
SELECTION (Y≤60)∧(B79MOD[;2]=13)

X LABEL
Y LABEL
NO. OF EL

CATEG.	NO.PTS	%-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	196	1	27.173	18.159	10	30	40
1	83	0.42347	29.108	17.674	12	36	40
3	56	0.28571	23.857	17.929	7	18	38
4	17	0.086735	28.412	21.992	3	25	49
5	19	0.096939	27.528	18.894	12	29	38
6	7	0.035714	29.857	11.532	15	35	40
7	14	0.071429	25.643	19.086	8	24	46

Figure 3.16 Equivalency Status vs. Diploma Grads

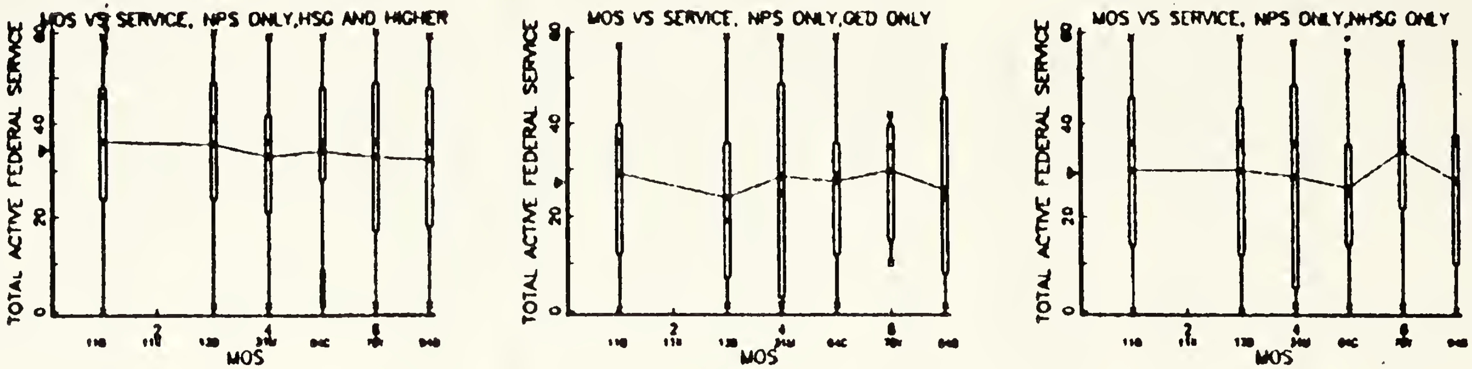


Figure 3.17 Recap of Effects of Education on Performance

2. Marital Status and Number of Dependents vs. Length of Service: (See Figures H.10 - H.15, Appendix H)
   - Most enlistees were single with no dependents at time of entry (92%)
   - Single enlistees with 1 dependent had highest mean performance.
   - In general, married enlistees outperformed single enlistees.
   - As number of dependents in married soldiers at time of entry increased, so did performance. (note: true up to 3 dependents.)
   - High-school-diploma graduates led in performance, followed by non-high-school-graduates, then GED holders.
3. Age Versus Length of Service: (See Figures H.16 through H.24, Appendix H)
   - As age increases (at time of entry) from 17 to 19 years, total service(performance) increases, followed by a leveling off in the 19 to 24 year range. Performance decreases as age increases from 24 to 29 years old.
   - High-school-diploma graduates outperformed other less educated entrants, non-high-school graduates followed, then equivalency certificate holders.
4. Sex Versus Length of Service: (See Figures H.25 through H.30, Appendix H)
   - Males outperformed females.
   - Diploma graduates (HSDG) outperformed all other education levels; Non-high-school grads outperformed the GED holders.
   - Non-high-school-graduate females outperformed male non-high-school-graduates and female high-school diploma graduates. However, non-high-school-graduate females displayed the greatest variance in

length of service, indicating a possible higher risk in attrition.

- Non-high-school-graduate females were followed by GED males and then NHSG males in variance in length of service (risk).

5. Race Versus Length of Service: (See Figures H.31 through H.36, Appendix H)
   - In general, the "other" category were the highest performers, followed by blacks, and then whites.
   - High-school-diploma graduates were the highest performers in terms of length of service in all categories.
   - Non-high-school-graduates outperformed the graduate equivalency certificate holders (GED).
   - GED blacks displayed the highest variance in length of service, indicating a higher risk in attrition, followed by GED whites.
6. Reenlistment Code Versus Education Level: (See Figures H.37 through H.39, Appendix H)
   - At least through the high school diploma graduate level of education, as education increased, reenlistment eligibility increased.
   - GED soldiers received about the same reenlistment codes as those soldiers who entered the army with 2 years of high school.
   - A wide variance in reenlistment eligibility was observed, possibly related to the fact that reenlistment eligibility is mostly up to the local commander's discretion.
7. Reenlistment Code Versus Length of Service: (See Figures H.39 through H.44, Appendix H)
   - Approximately 30% of the sample was uncoded at the time of the last file update, indicating that the bar to reenlistment may not used as often as it

could be as a rehabilitative tool for substandard soldiers.

- Generally, as length of service increased, so did the number of soldiers eligible to reenlist.
- Approximately 50% of the GED holders and of the non-high-school-graduates indicated a reenlistment code of 3, corresponding to only 30% in the high-school-diploma graduates. This code corresponds to a Department of the Army initiated bar to reenlistment. This proportion seems a bit unreasonable and more research seems necessary in this multivariate combination.
- High-school-diploma graduates outperformed all others, followed by non-high-school-graduates and the graduate equivalency certificate holders.
- Again, the GED holders demonstrated the largest variance (risk) in length of service, followed by non-high-school graduates.

8. <u>Character of Service Versus Education Level</u>: (See Figures H.45 through H.47, Appendix H)
   - Generally, as education level increased (at time of entry) through high school diploma, so did the character of service awarded.
   - GED and non-high-school-graduate soldiers received about the same treatment in character of service awarded. Again, high-school-diploma graduates received the largest proportion of the honorable discharges.
   - College graduates outperformed the high-school-diploma grads; however those entrants with one and two years of college were outperformed in terms of character of service awarded by the non-high school diploma graduates.

- A wide variance was noted possibly because of the commanders' discretion allowed in determining what character of service will be awarded to an individual.

9. <u>Character of Service Versus Length of Service</u>: (See Figures H.48 through H.54, Appendix H)
   - As length of service increased so did the number of honorable discharges awarded.
   - Over half of the cohort received an honorable discharge. A large proportion of early "leavers" also received this honorable discharge, perhaps indicating that some commanders are very lenient in their determination of type discharge to be awarded.
   - High-school-diploma graduates receiving honorable discharges performed the best in terms of length of service, again followed by non-high-school graduates and the GED's.
   - A wider variance in length of service was noted for the non-high-school graduates and the GED's who received honorable discharges, when compared to the high-school-diploma graduates. This perhaps relates again to the discretion that is exercised by the local commanders in awarding discharges. It seems that length of service may not be considered as an indicator of "good" service by a number of commanders in the field.

D. EXAMPLE CONFIRMATORY ANALYSIS

Conclusions drawn from the above explanatory data analysis must be analyzed in a formal manner to determine if differences, say, in the mean performance among the varying levels of education at entry is statistically significant.

An example is presented below; this type of analysis should be performed on all conclusions reached before any policy implementation.

The one-way-analysis of variance provides a well-structured approach in testing the equality of means among k sample populations. In this approach, the k populations are assumed

1. to be i.i.d normal populations, and
2. to have equal variances.

As has been pointed out, some of the boxplot analyses above has indicated that the equal variance assumption is obviously not true; contrarily, this difference in variance was was used as an indicator of the "risk" involved in recruiting that entrant with his particular qualifications. In those cases, nonparametric tests are available for confirmatory analysis. An example of that type will also be presented below. See [Ref. 15: pp. 492-503] for a discussion on 1-way ANOVA.

The actual calculations of the ANOVA table were completed using an APL program contained in public domain of the Naval Postgraduate School computer system (Library 5,OA3660). A copy of the program is at Appendix I. Results of the analysis are summarized in Table V Thus the null hypothesis can be rejected at the .05 level of significance.

In instances where the homogeneity of variance and normality of population assumptions are infeasible, nonparametric tests can be utilized for confirmation of statistical significance. The Kruskal-Wallis test was utilized to test the following hypothesis:

H0:mean services across all education levels are equal.

H1: for at least one pair of the population represented by the levels of education, the means are different. See [Ref. 16: pp.229-237] for a discussion of the K-W test.

TABLE V

Ancva Table for Testing Equality of Means

Z←ANOVA MATR

ANOVA WAS UPDATED 1/3/79, SEE ANOVAHOW FOR CHANGES.

ANOVA TABLE

SOURCE	DF	SS	MS	F
TREATMENT	4	25933.42	6483.35	21.67
ERROR	2998	896813.33	299.14	
TOTAL	3002	922746.75		

R-SQUARE = 0.028

OVERALL MEAN = 31.93

TREATMENT EFFECTS ¯4.76 2.69 ¯1.29 ¯3.24¯12.73

The actual calculations were performed utilizing the BMDP statistical scftware program P3S on the Naval Postgraduate School Computer System. See [Ref. 17: pp.442-443] for a description cf this package. Results are summarized in Table VI .

TABLE VI

Results of K-W Test of Equality of Means

Variable 1 LOS

Group No.	Name	Frequency	Rank Sum
1	SOPH	10	9064.0
2	JR	752	1017353.5
3	GED	196	250661.0
4	NHSG	504	730567.5
5	HSDG	1546	2517723.0

Kruskal-Wallis Test Statistic = 75.13
Level of Significance = 0.0000
Using Chi-Square Distributicn with 4 Degrees of Freedom

Thus the null hypothesis can be rejected at the .05 level of significance. Multiple comparisons may be calculated in accordance with [Ref. 16: p.231] to determine which pairs of means are different if desired.

E. SUMMARY OF EXPLORATORY DATA ANALYSIS EFFORTS

1. General

Exploratory Data Analysis techniques have been utilized in the forms of draftsman's displays and boxplots to "preprocess" a large data set. This analysis was presented as a demonstration of the power of EDA techniques and to provide initial insights into the attrition of U.S. Army enlistees prior to further analysis. These techniques have provided the following:

1. Familiarity with data set.
2. Reduction in the dimensionality of the data set; numerous variables were determined as having no appreciable effects on the dependent variable under consideration.
3. Identification of erroneous data.
4. Structure of the data set and variable coding.
5. Information on multivariate and pairwise associations among the variables. This information will be summarized below.
6. An intuitively pleasing form of analysis to assist analysts and decision makers in understanding the problem at hand.
7. A means of allowing the analyst to "compose" his method of attack on a large problem in an interactive fashion.

2. EDA and U.S. Army Enlisted Attrition

More specifically, in the investigation of attrition of U.S. Army FY79 enlistees, the following information has been revealed:

1. The variables listed in Table VII below were observed to have some effect on a soldiers performance which has been defined as his total length of service.

TABLE VII

Possible Explanatory Variables from EDA

Age
Sex
Race
Mental Category
Marital Status/Number of Dependents (at entry)
Education Level
Military Occupational Skill

2. Level of education at entry has an effect on the performance of enlistees. Education level seems to interact with the other variables listed in Table VII above, producing different levels of performance. Other insights provided by the analysis are
   - High school diploma graduates demonstrated better performance than did non-high-school-diploma graduates and those enlistees who had obtained a graduate equivalency certificate prior to entry.
   - Non-high school diploma graduates demonstrated better performance than did GED holders.
   - GED holders demonstrated a larger variance in total service obtained, indicating a higher risk in attrition than did non-high school diploma graduates and high school diploma graduates.

3. Character of Service and Reenlistment Code are two indicators of performance that may also be affected by education level at entry; however, the discretion exercised by local commanders in awarding these may have confounded them as suitable explanatory variables.
4. Confirmatory analysis needs to be performed after exploratory data analysis prior to forming final conclusion for policy implementation. An example was provided.

3. Limitations

The EDA techniques are not to be used in lieu of more classical statistical analysis; they are to be used in conjunction with them. Some limitations are

1. Acceptance of their use by other statisticians.
2. Package utilized in this thesis is in the experimental stage; others may not be readily available to the analyst. Since this package is experimental, certain capabilities are still being developed. For example, in the the ANCVA presented, the means used in the analysis were not stored in a global sense for further analysis and had to be entered "by hand" into an ANOVA program.
3. Cost of graphics capabilities for computer systems.
4. Storage necessary for EDA packages is currently not available for most personal or desk top computers.

Thus Exploratory Data Analysis has been shown to be a useful technique in the initial analysis of data. In the next chapters, a more formal analysis will be presented.

IV. SURVIVOR FUNCTION ANALYSIS

A. BACKGROUND

The Exploratory Data Analysis techniques presented (along with the necessary confirmatory analysis) have shown that level of education, sex, race, age, mental category, and marital status/number of dependents are candidate explanatory variables in determining the total length of service for an enlistee.

A survivor function approach is utilized to gain further insight into these explanatory and their relationship to length of service.

Suppose that the length of service of an enlistee is a random variable X. The cumulative distribution function or c.d.f., then, can be viewed as giving the probability that an enlistee will "die" or "fail" or leave the army before x units of time, a realization of the random variable X, have elapsed. Then the quantity

$$S(x) = 1-F(x) = P(X>x) \qquad \text{(eqn 4.1)}$$

called the survivor function, provides the probability that an enlistee "survives" more than x units of time. The survivor function, for discrete data, is a step function, where the height of the jump between any two values of x is equal to $P(X=x)$. The survivor function is estimated by using the following relative frequency definition of probability:

$$S(x) = 1-F(x) = (\text{number of observations} > x)/n \qquad \text{(eqn 4.2)}$$

where n is the number in the sample being considered [Ref. 18: pp. 92-93, 263-264].

B. APPLICATION

The survivor function is therefore a logical means of analyzing enlistee behavior with regards to length of service. For this portion of the analysis, attrition will be defined as failure to complete the first term of service. The FY79 COHORT consists of three-year obligated enlistees (3YO) and four-year obligated enlistees (4YO). Each subset will be analyzed separately. The reenlistment decision will be defined as completing greater than one term of service. Hence, based on the survival function model, the following equations indicate the "life" cycle of the enlistee:

TABLE VIII

Enlistee Life Cycle Models

<u>3 Year Enlistees</u>

P(enlistee will attrite) = P(X≤36 mos.)
= 1-P(X>36 mos.) = 1-S(x)

P(enlistee will complete 1 term) = P(X=36 mos.)
= Ht. of jump at x=36

P(enlistee will reenlist) = P(X>36 mos.) = 1-S(x)

<u>4 Year Enlistees</u>

P(enlistee will attrite) = P(X≤48 mos.)
= 1-P(X>48 mos.) = 1-S(x)

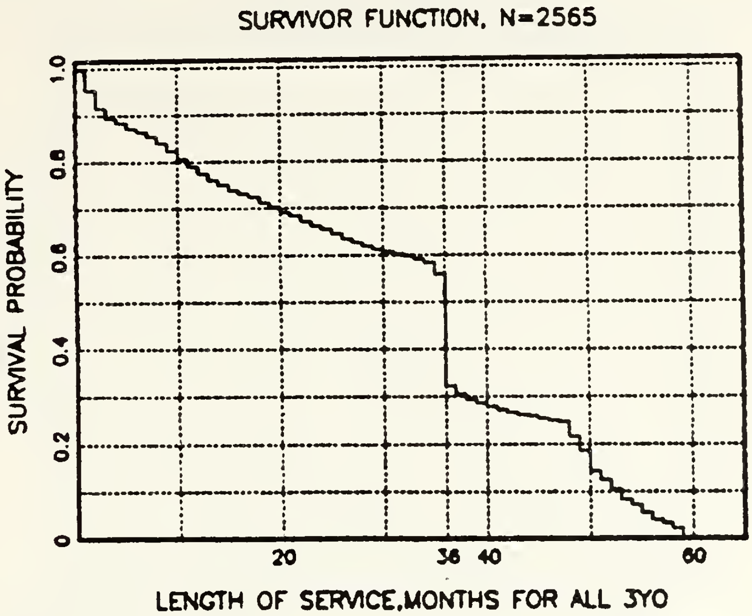
P(enlistee will complete 1 term) = P(X=48 mos.)
= Ht. of jump at x=48

P(enlistee will reenlist) = P(X>48 mos.) = 1-S(x)

These realization times for the random variable X are chosen just to demonstrate the survivor function methodology. Considerable research has been done in defining the exact time cutoff for the definition of attrition, since the first "term" of service may actually be thirty-four months as opposed to thirty-six because of the various "early out" programs offered by different commands. This analysis does allow for at least this concept in its definition of attrition as strictly less than thirty-six and forty-eight months (i.e. thirty-five and forty-seven months) length of service for the three year and four year obligations respectively. The methodology suffices no matter where the attrition definition is made in the length of service parameter.

Survivor functions for each of the previously defined candidate explanatory variables are estimated utilizing the Cumulative Distribution input screen in the IBM GRAFSTAT data analysis package. (See Appendix J for a depiction of this screen.) These functions are analyzed for the above mentioned statistics.

The survivor function of the entire sample across all variables is presented to demonstrate the characteristics of the analysis in Figure 4.1 . Using this type of analysis, statistics for the most prevalent enlistee are presented in Figure 4.2 and in Figure 4.3 for 3 year enlistees and 4 year enlistees respectively. The "most prevalent enlistee" was determined by observing the total number in each of the seven variables being considered. Note that the four year obligated enlistees demonstrated higher probability of attrition (0.33 to 0.24), and a lower probability of reenlistment (0.12 to 0.37) than the three year obligated enlistee, ceteris paribus. This may indicate that the utilization of the three year term of service is more cost effective than the four year term, considering "assess,dress,train" cost described earlier. Results for the most prevalent enlistee are summarized in Table IX .



SAMPLE GRAPHICAL CALCULATIONS

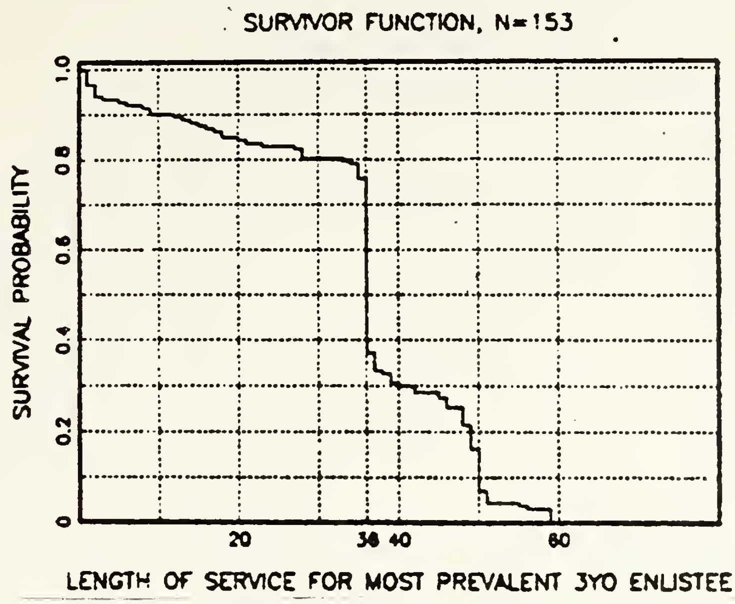
P(LENGTH OF SERVICE=36MONTHS)
=HEIGHT OF JUMP AT LOS=36 MONTHS
0.559−0.320=0.239

P(LENGTH OF SERVICE<36MONTHS)
=VERTICAL DROP IN CURVE FROM 0 TO 35 MONTHS
1.0−0.559=0.441

P(LENGTH OF SERVICE>36MONTHS)
=VERTICAL DROP IN CURVE FROM 37 TO 59 MONTHS
0.321−0.000=0.321

Figure 4.1 Survivor Function For All 3 YO Enlistees

The effects of education level on attrition of three-year-obligated enlistees are presented in Figure 4.4 . Those enlistees with a graduate equivalency certificate indicate a higher probability of attrition (0.54) than both those with two years of high school and those with non-high-school-diploma graduate (NHSDG) status, having three to four years of high school. These findings reinforce the earlier boxplot analysis of length of service. The trend is also

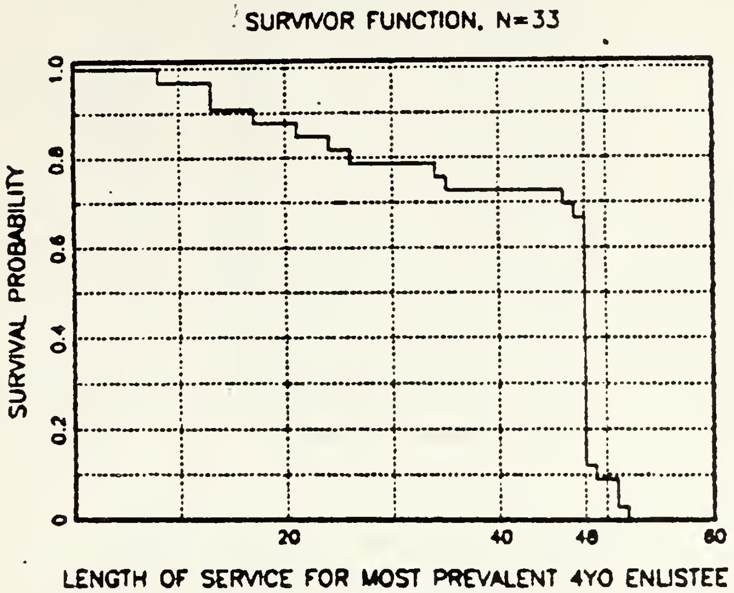


THIS BATCH CONSISTS OF THE MOST PREVALENT ENLISTEES: 18 YR OLD, COMBAT ARMS,3YO, SINGLE WITH NO DEPS, WHITE MALE

P(LOS<36)=0.242
P(LOS=36)=0.385
P(LOS>36)=0.373
AVERAGE LOS=35.529

Figure 4.2 Survivor Function, Most Prevalent 3YO Enlistee

evident in the probability of reenlistment with the high-school-diploma graduates having the highest, followed by the NHSDG, then those with 2 years high school and then those enlistees with equivalency certificates. Again the GED enlistee is seen to be inferior to the non high school diploma graduate (NHSDG). Results of this survivor function analysis are in Table X in the following section.



THIS BATCH CONSISTS OF THE MOST PREVALENT 4YO ENLISTEE: 18 YR OLD COMBAT ARMS, SINGLE WITH NO DEPS,WHITE,MALE

P(LOS<48)=0.334
P(LOS=48)=0.545
P(LOS>48)=0.121
AVERAGE LOS=40.939

Figure 4.3 Survivor Function, Most Prevalent 4YO Enlistee

TABLE IX

Results of Survivor Analysis on Most Prevalent Enlistee

Term	P(attrite)	P(full term)	P(Reenlist)	Ave LOS
3 Years	0.242	0.385	0.373	35.53
4 Years	0.334	0.545	0.121	40.94

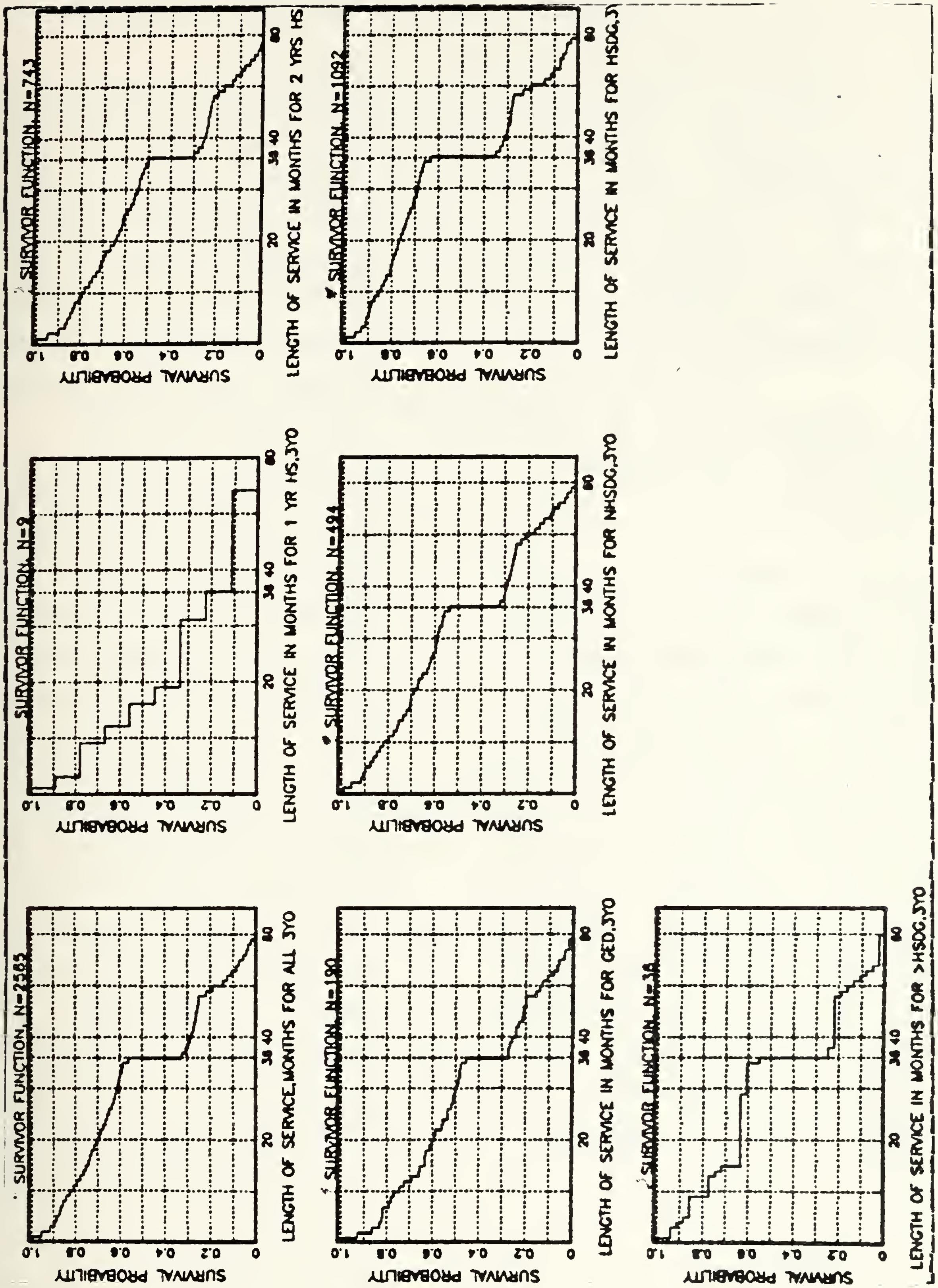


Figure 4.4 Survivor Functions of Education Levels

C. RESULTS OF SURVIVOR FUNCTION ANALYSIS

Similar analyses were performed on the remaining six candidate explanatory variables for both the three year enlistees and for the four year enlistees. Actual survivor functions are found in Appendix K and L respectively. Tabular summaries are provided in Table X and XI . Based on these summary tables, the lowest risk attributes in terms of lowest probability of attrition are provided in Table XII . Those attributes that result in the highest probability of reenlistment are shown in Table XIV and Table XV .

The usefulness and simplicity of this approach is appealing for "quick and dirty" analyses in determining short term policy decisions concerning the characteristics desired for prospective soldiers.

Modeling of the survivor curves can provide further prediction capabilities and significance of each of the covariates. However, the survivor functions presented all demonstrate the large "jump" at thirty-six and forty-eight months for the three year and four year enlistees, respectively. This large "discontinuity" causes modelling of the survivor function to be somewhat difficult. Modelling of the survivor is discussed in great detail in [Ref. 19] and [Ref. 20]. The noted similarity in the survivor curves for all the variables under investigation suggest that the use of a Cox proportional-hazards model may be appropriate; however more research is needed in the modeling of the discrete jump in the survivor curve. See [Ref. 20]. This analysis has been presented as an initial methodology to demonstrate its usefulness ; no modelling will be performed.

TABLE X

Results of Survivor Function Analysis for 3YO

Variable	P(LOS<36)	P(LOS=36)	P(LOS>36)	Ave. LOS
All variables	0.441	0.239	0.321	30.87
Education lvl				
1 yr High Sch	0.778	0.111	0.111	20.11
2 yrs HS	0.503	0.201	0.296	28.72
GED	0.542	0.184	0.274	27.09
NHSDG	0.461	0.217	0.322	30.69
HSDG	0.366	0.283	0.351	33.23
College	0.400	0.400	0.200	32.80
Sex				
Male	0.431	0.246	0.322	31.17
Female	0.515	0.175	0.310	28.39
Race				
White	0.467	0.258	0.275	29.21
Black	0.405	0.204	0.391	33.18
Other	0.374	0.238	0.388	34.02
Mental Cat				
Cat1	0.357	0.214	0.429	33.57
Cat2	0.401	0.294	0.305	31.85
Cat3a	0.356	0.320	0.324	32.40
Cat3b	0.457	0.225	0.318	30.42
Cat4a	0.449	0.230	0.321	29.86
Cat4b	0.468	0.222	0.310	30.80
Cat4c	0.443	0.193	0.364	32.57
Mar.Stat/No.Deps				
Single/0	0.441	0.252	0.307	30.62
Married/0	0.448	0.083	0.469	32.92
Married/1	0.563	0.000	0.437	30.19
Married/2	0.381	0.095	0.524	36.76
Married/3	0.333	0.083	0.583	42.00
Military Skill				
11B	0.461	0.225	0.314	30.28
13B	0.464	0.195	0.341	30.73
31M	0.408	0.241	0.351	31.07
64C	0.419	0.302	0.279	31.18
76Y	0.402	0.217	0.381	32.82
94B	0.428	0.260	0.312	30.85
Age				
17	0.491	0.215	0.294	29.65
18	0.432	0.255	0.313	31.04
19	0.396	0.266	0.338	32.09
20	0.482	0.232	0.286	28.92
21	0.405	0.250	0.345	31.61
22	0.495	0.206	0.299	28.80
23	0.446	0.149	0.405	33.15
24	0.450	0.225	0.325	32.18
25	0.470	0.118	0.412	32.77
26	0.556	0.074	0.370	25.04
27	0.450	0.150	0.400	30.80
28	0.364	0.182	0.454	31.10
29	0.571	0.143	0.286	27.00
30	0.286	0.000	0.714	53.14
> 30	0.524	0.238	0.238	25.86

TABLE XI

Results of Survivor Function Analysis for 4YO

Variable	P(LOS<48)	P(LOS=48)	P(LOS>48)	Ave. LOS
All variables	0.441	0.239	0.321	30.87
Education lvl				
1 yr High Sch	-	-	-	-
2 yrs HS	0.670	0.110	0.230	31.78
GED	0.666	0.167	0.167	29.67
NHSDG	0.500	0.125	0.375	35.88
HSDG	0.367	0.409	0.223	39.18
College	0.600	0.240	0.160	28.96
Sex				
Male	0.391	0.389	0.220	38.29
Female	0.667	0.000	0.333	37.00
Race				
White	0.387	0.389	0.220	37.74
Black	0.424	0.288	0.288	39.36
Other	0.307	0.500	0.193	39.42
Mental Cat				
Cat1	0.333	0.429	0.238	41.14
Cat2	0.376	0.368	0.256	38.68
Cat3a	0.324	0.485	0.191	39.78
Cat3b	0.427	0.427	0.146	37.29
Cat4a	0.394	0.346	0.260	38.10
Cat4b	0.465	0.296	0.239	36.66
Cat4c	--	--	--	--
Mar.Stat/No.Deps				
Single/0	0.393	0.402	0.205	37.88
Married/0	0.550	0.150	0.300	38.85
Married/1	--	--	--	--
Married/2	--	--	--	--
Military Skill				
11B	0.359	0.412	0.229	39.43
13B	0.434	0.352	0.214	36.73
31M	--	--	--	--
64C	--	--	--	--
76Y	--	--	--	--
94B	--	--	--	--
Age				
17	0.418	0.436	0.145	36.86
18	0.349	0.429	0.222	39.56
19	0.353	0.431	0.216	39.12
20	0.797	0.018	0.185	28.92
21	0.455	0.394	0.151	36.73
22	0.476	0.238	0.286	38.19
23	0.662	0.040	0.298	33.15
24	0.675	0.050	0.275	32.18
25	0.647	0.000	0.353	32.77
26	0.500	0.125	0.375	33.50
28	--	--	--	--
29	--	--	--	--
30	--	--	--	--
> 30	--	--	--	--

TABLE XII

Attribute Values Providing Lowest Risk of Attrition,3YO

Variable	Lowest P(attrition)	Next lowest P(attrition)
Education Level	HSDG	College
Sex	Male	Female
Race	Other	Black
Mental Cat.	3A	1
Marital Stat/ No. of Deps	M/2	S/0
Age	28	19
MCS	76Y	31M

TABLE XIII

Attribute Values Providing Lowest Risk of Attrition,4YO

Variable	Lowest P(attrition)	Next lowest P(attrition)
Education Level	HSDG	NHSDG
Sex	Male	Female
Race	Other	White
Mental Cat.	3A	1
Marital Stat/ No. of Deps	S/0	M/0
Age	18	19
MCS	11E	31M

TABLE XIV

Attribute Values Providing Highest Reenlistment,3 YO

Variable	Lowest P(reenlist)	Next lowest P(reenlist)
Education Level	HSDG	NHSDG
Sex	Male	Female
Race	Black	Other
Mental Cat.	1	4C
Marital Stat/ No. of Deps	M/3	M/2
Age	30	28
MCS	76Y	31M

TABLE XV

Attribute Values Providing Highest Reenlistment,4YO

Variable	Lowest P(reenlist)	Next lowest P(reenlist)
Education Level	NHSDG	2 yrs HS
Sex	Female	Male
Race	Black	White
Mental Cat.	4A	2
Marital Stat/ No. of Deps	M/0	S/0
Age	26	25
MCS	11B	13B

V. RESULTS AND CONCLUSIONS

A. GENERAL

An intuitively pleasing,simple methodology was presented for the study of performance in the form of length of service of U.S Army enlistees. Some Exploratory Data Analysis techniques were demonstrated through the use of the IBM GRAFSTAT data analysis package. The interactive capabilities of the package and the APL language were explcited to provide a means of rapidly manipulating and observing the selected data. The tools proposed, the draftsman's displays, boxplots and survivor functions, were used on actual cohort data from the Defense Manpower Data Center.

Several possible explanatory variables and their association with performance were presented based on the Exploratory Data Analysis. Confirmatory analysis was performed to support the Exploratory Data Analysis. Probabilities of enlistee attrition and reenlistment were provided using a survivor function analysis for each of the candidate explanatory variables. Attributes that presented the highest risk of attrition and the highest probability were presented.

B. SUMMARY

The increasing cost of "assessing,dressing and training" today's Army enlistee coupled with the diminishing supply of 17-21 year old prospective enlistees have prompted research effort toward gaining insight into those personal attributes that produce the most successful soldier in terms of first term completion. The basis for understanding the relationships of these personal attributes and for using this

understanding in recruiting policy lies in the ability to rapidly analyze the available data on current enlistees and to present the analysis in a form that is understandable and useful for the decision maker.

This thesis has presented a broadly applicable and simple methodology, using Exploratory Data Analysis through the interactive capability of the IBM GRAFSTAT package and the APL language, for defining the area of analysis, identifying errors in the data, reducing the dimensionality of the problem, and determining relevant association of personal characteristics of enlistees to performance. The bonds between Exploratory Data Analysis and more classical statistical analysis were demonstrated. Use of survivor function analysis provided statistics on chosen explanatory variables, indicating the importance of these characteristics.

The further application of the methods in this thesis and of Exploratory Data Analysis in general should increase the practitioner's ability to make sound decisions regarding future manpower planning issues. With the increased availability of graphics-capable personal computers, Exploratory Data Analysis is relevant at all levels of decision making.

C. RECOMMENDED FURTHER RESEARCH

The following items deem further research:

1. A comparison of Exploratory Analysis techniques in this theses to other data analysis packages such as those available in BMDP and SAS (see [Ref. 17] and [Ref. 21]) would be useful in determining advantages and disadvantages of the different approaches in variable selection and error identification. In particular, the Cox proportional hazards model in the BMDP program P2L [Ref. 17: pp. 576-594] uses a stepwise approach to identify important explanatory

variables and estimates the survivor as well as the hazard function for further analysis. Note however that models such as this Cox proportional hazards model for estimating the effect of concomitants on survival curves are not applicable because of jumps in the survival functions at known times. More research is needed to determine how to apply the model to such a function with this large discrete jump.

2. The Graduate Equivalency Degree programs offered throughout the United States need to be analyzed in detail for standards used in awarding the certificates. The wide variance and the poorer performance of the GED holders indicate that non-high-school-diploma graduates should be treated separately in any analysis, contrary to the popular grouping of the two categories. Perhaps different GED levels would provide insight into future performance at least as well as the different levels of high school status have.
3. The trends and probabilities have been presented as a methodology on only a 10 percent sample of the data: comparisons of these outcomes to other data sets would be useful in the determination of prediction possibilities.
4. Modeling of the survivor curves would provide a detailed account of the actual contribution of each explanatory variables using multivariate regression techniques. Again further research is needed in the applications of modeling techniques to survivor curves with the noticeable jumps at known times.

APPENDIX A
EXPLORATORY DATA ANALYSIS TECHNIQUES

Exploratory Data Analysis techniques are usually first attributed to John W. Tukey in his book by that title [Ref. 22]. Exploratory Data Analysis for the purposes of this thesis will be defined as "the activity of examining data, both graphically and through numerical summaries, for the purpose of revealing properties of the data itself, and with luck, of the processes giving rise to that data." [Ref. 23: p. 2]. Thus EDA techniques can be thought of as "informal " techniques to examine the data prior to "formal", more classical analysis techniques, in order to prevent needless calculations irrelevant to the investigation at hand. Quite often more can be learned about the data in this initial, informal look at the data. As Chambers et. al. points out, graphical EDA methods are perhaps most effective in the initial glance at the data to limit the scope of the investigation to only those variables that are pertinent [Ref. 12]. These graphical methods allow the investigator to rapidly synthesize information, in a more efficient and intuitive manner perhaps than through methods available in commercial statistical packages that produce tabular data.

One particular method of multivariate analysis is the multidimensional array of scatter plots called a "generalized draftsman's display"of the data [Ref. 12: pp. 136,145]. An example display is seen in Figure A.1 This figure demonstrates how the pairwise scatter plots are arranged so that "any adjacent pair of plots have an axis in common" [Ref. 12: p. 145]. All variables of interest, then, for the entire data set can be displayed as the first phase of

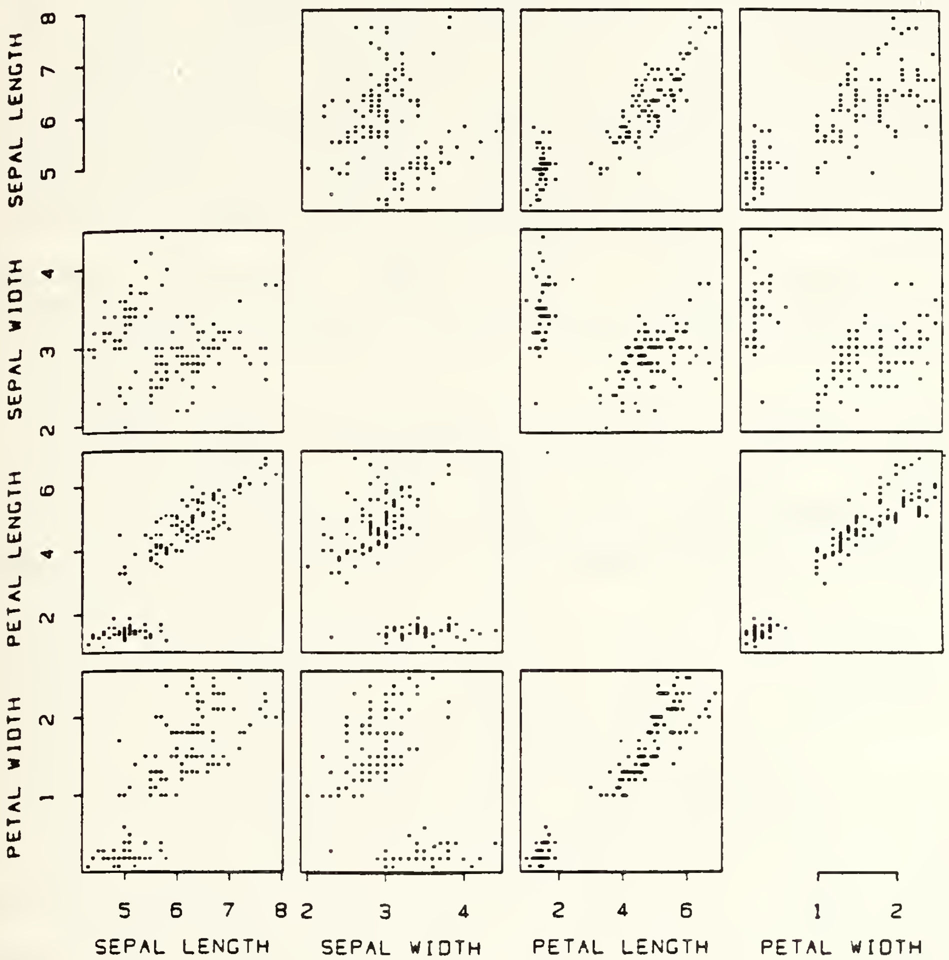


Figure A.1 Draftsman Display from Chambers et. al.

the investigation. Then one can rapidly and effectively determine if trends exist and for which specific pairwise association of variables.

Captain Malcolm Johnson, a student of Operations Research at the Naval Postgraduate School, has developed an APL program called "draftsman" that is imbedded in IBM's GRAFSTAT package on the school's computer system (See [Ref. 24: pp. 13-17]. for information on use of GRAFSTAT) that organizes any data set into a draftsman's display. This program also allows for transformations of the data and for jittering of the data. His efforts have been published as a Master's thesis that includes a tutorial for use of the "draftsman" program. [Ref. 13]. This program will be utilized in the initial phase of the data analysis efforts of this thesis.

Of course, the draftsman's display is only the first step of the analysis. If trends are evident, then further analysis should be performed utilizing more formal confirmatory analysis to verify any graphically-determined associations among the variables.

The use of boxplots is another EDA technique that is very useful in "taking an initial look" at the data. The boxplot is a "simple method of summarization".

The upper and lower quartiles are depicted by the "body" of the box, the median is portrayed by a line, circle or other distinguishing mark as is the mean. Upper and lower adjacents are depicted at the end of lines extending from the body of the box. These terms are defined as the "largest observation that is less than or equal to the upper quartile plus 1.5 times the interquartile range, and smallest observation that is greater than or equal to the lower quartile minus 1.5 times the interquartile range," respectively. Values that fall outside the range of adjacent values are called outside values. These are plotted as

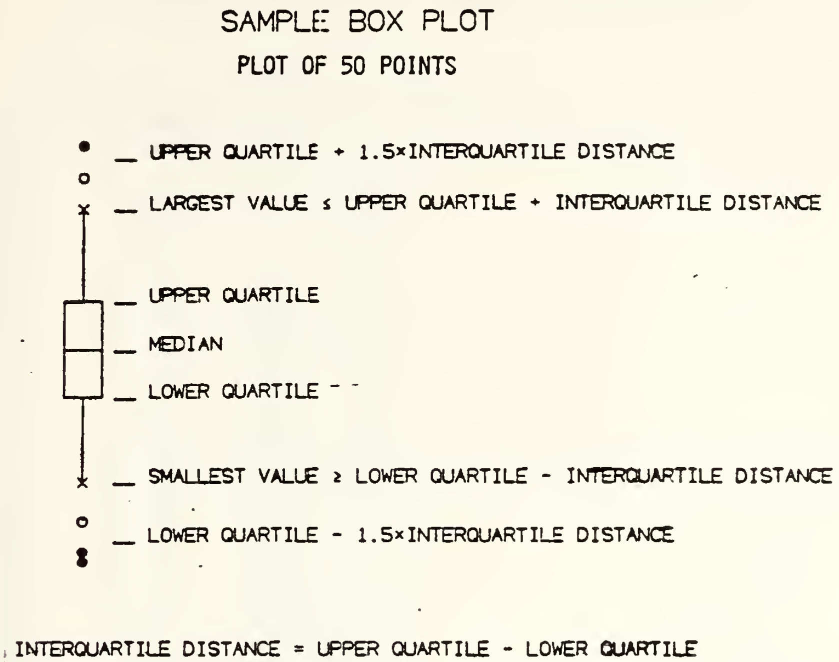
individual points. See Figure A.2 for a depiction of the components of the boxplot.

The boxplot provides a rapid "impression" of the distribution of the data. The median, mean, and spread are all obvious. The length of the lines to the adjacent values demonstrate the "stretch" of the tails of the distribution. The individual points for the outside values allows the user of the plot to consider "outliers" although not every outside value is an outlier.

The figure also allows for some determination of the symmetry of the distribution of the data, simply by viewing the symmetry of the body of the box about the median line or dot.

These plots are useful when it is not feasible or necessary to capture all the details of a distribution, or when many distributions need be compared. The width of the box has no significance.

An excellent discussion of these and other EDA techniques is found in [Ref. 12], from which this description of boxplots was taken.



Example box plot for fifty data points from a regenerative simulation. The interquartile distance equals the estimated upper quartile minus the estimated lower quartile. The light circles are data points which fall between the largest value less than or equal to the upper quartile plus the interquartile distance and the upper quartile plus 1.5 times the interquartile distance. The dark circles are data with values above this latter point. Similarly for the lower part of the box plot.

Figure A.2 Example Boxplot

APPENDIX B

HISTOGRAM OF FY79 NHSG ASSESSIONS

TABLE NO. 1
MOS DISTRIBUTION FOR ARMY NPS ACCESSIONS.

CODE	FREQUENCY	PERCENT	CUM. FREQ.	CUM. PCNT.	0....5....10....15....20....25....30....
AR	1	0.0	1	0.0	
A6J	1	0.0	2	0.0	
CMF	1	0.0	3	0.0	
C21	1	0.0	4	0.0	
C5B	2	0.0	6	0.0	
CFS	1	0.0	7	0.1	
C13	1	0.0	8	0.1	
C28	1	0.0	9	0.1	
C2F	2	0.0	11	0.1	
C3C	4	0.1	15	0.1	
C5R	99	0.8	117	0.9	*
05C	233	1.8	350	2.8	**
05D	1	0.0	351	2.8	
05K	1	0.0	352	2.8	
05L	1	0.0	353	2.8	
08G	1	0.0	354	2.8	
11B	4	0.0	358	2.8	
11C	1	0.0	359	2.8	
11E	1	0.0	360	2.8	
11T	1	0.0	361	2.8	
11X	2258	17.8	2619	20.6	******************
11Y	5	0.0	2624	20.7	
12B	348	2.7	2972	23.4	***
12C	60	0.5	3032	23.9	
12E	38	0.3	3070	24.2	
12F	21	0.2	3091	24.3	
13B	536	4.2	3627	28.5	****
13C	9	0.1	3636	28.6	
13E	117	0.9	3753	29.5	*
13F	113	0.9	3866	30.4	*
13G	1	0.0	3867	30.4	
13M	28	0.2	3895	30.7	
13R	11	0.1	3906	30.8	
13T	1	0.0	3909	30.8	
15D	74	0.6	3983	31.3	*
15E	90	0.7	4073	32.1	*
15G	1	0.0	4074	32.1	
15J	7	0.1	4081	32.1	
16B	26	0.2	4107	32.3	
16C	20	0.2	4127	32.5	
16D	20	0.2	4147	32.6	
16E	34	0.3	4181	32.9	
16H	3	0.0	4184	32.9	
16J	22	0.2	4206	33.1	
16K	1	0.0	4207	33.1	
16P	47	0.4	4254	33.5	
16R	126	1.0	4380	34.5	*
16S	84	0.7	4464	35.2	*
16X	1	0.0	4465	35.2	
17B	19	0.1	4484	35.3	
17C	19	0.1	4503	35.4	
17K	11	0.1	4514	35.5	
17P	1	0.0	4515	35.5	
19A	589	4.6	5104	40.2	*****
19C	2	0.0	5106	40.2	
19D	227	1.8	5333	42.0	**
19E	1	0.0	5334	42.0	
19F	2	0.0	5336	42.0	
19K	76	0.6	5412	42.6	*
21B	1	0.0	5413	42.6	
21L	14	0.1	5427	42.7	
22L	3	0.0	5430	42.7	
22N	8	0.1	5438	42.8	
22X	1	0.0	5439	42.8	
23A	2	0.0	5441	42.8	
24C	21	0.2	5462	43.0	
24G	5	0.0	5467	43.0	
24H	1	0.0	5468	43.0	
24J	2	0.0	5470	43.1	
24K	2	0.0	5472	43.1	
24L	2	0.0	5474	43.1	
24M	30	0.2	5504	43.3	
24N	20	0.2	5524	43.5	
24P	1	0.0	5525	43.5	
24Q	4	0.0	5529	43.5	

TABLE NO. 1
MOS DISTRIBUTION FOR ARMY NPS ACCESSIONS.

CODE	FREQ-UENCY	PER-CENT	CUM. FREQ.	CUM. PCNT.	0....5....10....15....20....25....30....35....40....
24U	15	0.1	5544	43.6	
24V	1	0.0	5545	43.6	
25L	9	0.1	5554	43.7	
26B	2	0.0	5556	43.7	
26C	5	0.0	5561	43.8	
26D	1	0.0	5562	43.8	
26G	1	0.0	5563	43.8	
26L	10	0.1	5573	43.9	
26Q	36	0.3	5609	44.1	
26R	4	0.0	5613	44.2	
26V	20	0.2	5633	44.3	
26Y	7	0.1	5640	44.4	
27B	3	0.0	5643	44.4	
27E	82	0.6	5725	45.1	*
27F	1	0.0	5726	45.1	
27G	5	0.0	5731	45.1	
27N	10	0.1	5741	45.2	
31B	1	0.0	5742	45.2	
31C	1	0.0	5743	45.2	
31D	1	0.0	5744	45.2	
31E	81	0.6	5825	45.8	*
31J	27	0.2	5852	46.1	
31M	387	3.0	6239	49.1	***
31N	31	0.2	6270	49.4	
31S	5	0.0	6275	49.4	
31T	5	0.0	6280	49.4	
31U	1	0.0	6281	49.4	
31V	146	1.1	6427	50.6	*
31W	1	0.0	6428	50.6	
31S	1	0.0	6429	50.6	
32D	22	0.2	6451	50.8	
32F	6	0.0	6457	50.8	
32G	1	0.0	6458	50.8	
32H	8	0.1	6466	50.9	
32J	1	0.0	6467	50.9	
33S	1	0.0	6468	50.9	
34Y	7	0.1	6475	51.0	
35B	1	0.0	6476	51.0	
35C	2	0.0	6478	51.0	
35D	1	0.0	6479	51.0	
35E	17	0.1	6496	51.1	
35F	3	0.0	6499	51.2	
35G	2	0.0	6501	51.2	
35H	3	0.0	6504	51.2	
35K	15	0.1	6519	51.3	
35L	1	0.0	6520	51.3	
36C	195	1.5	6715	52.8	**
36D	1	0.0	6716	52.8	
36E	6	0.0	6722	52.9	
36H	11	0.1	6733	53.0	
36K	217	1.7	6950	54.7	**
36R	1	0.0	6951	54.7	
41C	13	0.1	6964	54.8	
42D	1	0.0	6965	54.8	
43E	61	0.5	7026	55.3	
43M	1	0.0	7027	55.3	
44B	23	0.2	7057	55.5	
44E	5	0.0	7062	55.6	
45B	2	0.0	7064	55.6	
45D	62	0.5	7126	56.1	
45E	6	0.0	7132	56.1	
45G	8	0.1	7140	56.2	
45K	44	0.3	7184	56.5	
45L	20	0.2	7204	56.7	
45M	2	0.0	7206	56.7	
45N	22	0.2	7228	56.9	
45T	40	0.3	7268	57.2	
46N	10	0.1	7278	57.3	
50D	1	0.0	7279	57.3	
51B	30	0.2	7309	57.5	
51C	6	0.0	7315	57.6	
51K	38	0.3	7353	57.9	
51M	3	0.0	7356	57.9	
51N	11	0.1	7367	58.0	
51R	18	0.1	7385	58.1	

TABLE NO. 1

MOS DISTRIBUTION FOR ARMY NPS ACCESSIONS.

CODE	FREQ-UENCY	PER-CENT	CUM. FREQ.	CUM. PCNT.	0....5....10...15...20...25...30...35...40...
52C	43	0.3	7429	58.5	
52D	119	0.9	7547	59.4	*
54C	21	0.2	7568	59.6	
54E	135	1.1	7703	60.6	*
54G	1	0.0	7704	60.6	
55B	62	0.5	7766	61.1	
55D	9	0.1	7775	61.2	
55G	1	0.0	7776	61.2	
57E	8	0.1	7784	61.3	
57F	4	0.0	7788	61.3	
57H	17	0.1	7805	61.4	
6CH	1	0.0	7806	61.4	
6CS	1	0.0	7807	61.4	
61B	30	0.2	7837	61.7	
61C	27	0.2	7864	61.9	
61F	2	0.0	7866	61.9	
62B	111	0.9	7977	62.8	*
62C	19	0.1	7996	62.9	
62F	46	0.4	8042	63.3	
62G	1	0.0	8043	63.3	
62H	1	0.0	8044	63.3	
62J	30	0.2	8074	63.5	
62S	1	0.0	8075	63.6	
63B	440	3.5	8515	67.0	***
63C	3	0.0	8518	67.0	
63D	97	0.8	8615	67.8	*
63E	14	0.1	8629	67.9	
63G	26	0.2	8655	68.1	
63H	112	0.9	8767	69.0	*
63J	23	0.2	8790	69.2	
63M	1	0.0	8791	69.2	
63N	163	1.3	8954	70.5	*
63P	1	0.0	8955	70.5	
63S	86	0.7	9041	71.2	*
63T	213	1.7	9254	72.8	**
63W	62	0.5	9316	73.3	
63Y	50	0.4	9366	73.7	
63J	1	0.0	9367	73.7	
64B	1	0.0	9368	73.7	
64C	564	4.4	9932	78.2	****
64G	1	0.0	9933	78.2	
65B	1	0.0	9934	78.2	
67G	5	0.0	9939	78.2	
67H	4	0.0	9943	78.3	
67N	67	0.5	10010	78.8	*
67T	4	0.0	10014	78.8	
67U	29	0.2	10043	79.0	
67V	44	0.3	10087	79.4	
67X	1	0.0	10088	79.4	
67Y	36	0.3	10124	79.7	
68B	16	0.1	10140	79.8	
68D	16	0.1	10156	79.9	
68F	8	0.1	10164	80.0	
68G	50	0.4	10214	80.4	
68H	12	0.1	10226	80.5	
68J	32	0.3	10258	80.7	
68M	13	0.1	10271	80.8	
71C	1	0.0	10272	80.8	
71D	2	0.0	10274	80.8	
71G	9	0.1	10283	80.9	
71L	64	0.5	10347	81.4	*
71M	10	0.1	10357	81.5	
71N	6	0.0	10363	81.6	
71P	21	0.2	10384	81.7	
72E	58	0.5	10442	82.2	
72G	38	0.3	10480	82.5	
73C	10	0.1	10490	82.6	
73D	1	0.0	10491	82.6	
74C	2	0.0	10493	82.6	
74D	4	0.0	10497	82.6	
74F	2	0.0	10499	82.6	
75B	76	0.6	10575	83.2	*
75C	23	0.2	10598	83.4	
75D	83	0.7	10681	84.1	*
75E	29	0.2	10710	84.3	

TABLE NO. 1
MOS DISTRIBUTION FOR ARMY NPS ACCESSIONS.

CODE	FREQ-UENCY	PER-CENT	CUM. FREQ.	CUM. PCNT.	0....5....10....15....20....25....30....35....40...
75F	4	0.0	10714	84.3	
75J	1	0.0	10715	84.3	
76C	141	1.1	10856	85.4	*
76D	1	0.0	10857	85.5	
76J	13	0.1	10870	85.6	
76L	1	0.0	10871	85.6	
76P	40	0.3	10911	85.9	
76U	1	0.0	10912	85.9	
76V	86	0.7	10998	86.6	*
76W	91	0.7	11089	87.3	*
76X	2	0.0	11091	87.3	
76Y	253	2.0	11344	89.3	**
81B	5	0.0	11349	89.3	
81C	2	0.0	11351	89.3	
82C	109	0.9	11460	90.2	*
83E	1	0.0	11461	90.2	
83F	2	0.0	11463	90.2	
84B	1	0.0	11464	90.2	
84C	1	0.0	11465	90.2	
85B	1	0.0	11466	90.3	
91A	1	0.0	11467	90.3	
91B	327	2.6	11794	92.8	***
91C	11	0.1	11805	92.9	
91D	4	0.0	11809	92.9	
91E	24	0.2	11833	93.1	
91G	1	0.0	11834	93.1	
91H	4	0.0	11838	93.2	
91L	1	0.0	11839	93.2	
91R	16	0.1	11855	93.3	
91Y	1	0.0	11856	93.3	
92B	1	0.0	11857	93.3	
93C	1	0.0	11858	93.3	
93F	11	0.1	11869	93.4	
93H	3	0.0	11872	93.4	
93J	3	0.0	11875	93.5	
94B	398	3.1	12273	96.6	***
94E	1	0.0	12274	96.6	
94F	12	0.1	12286	96.7	
95B	387	3.0	12673	99.7	***
95C	26	0.2	12699	100.0	
96B	2	0.0	12701	100.0	
96C	1	0.0	12702	100.0	
96H	1	0.0	12703	100.0	
98J	2	0.0	12705	100.0	

APPENDIX C

FORTRAN PROGRAMS TO READ DATA

```
C
C
//DTHOMAS JOB (1972,0398),'D.THOMAS',CLASS=B
//*MAIN ORG=NPGVM1.1972P
//   EXEC FORTXCG
//FORT.SYSIN DD *
          INTEGER*4    SSAN1,SSAN2,SSAN3,
     *    DOSYC,DOSMO,DOSDO,DOSYS,DOSMS,DOSDS
          INTEGER*4 BLANK,NINE
          DATA BLANK/'    '/,NINE/' -9'/
       DO 300 I=1,30778,10
C
       SSAN1    = BLANK
       SSAN2    = BLANK
       SSAN3    = NINE
       DOSYC    = -9
       DOSMO    = -9
       DOSDO    = -9
       DOSYS    = -9
       DOSMS    = -9
       DOSDS    = -9.
C
          READ(1,100) SSAN1,SSAN2,SSAN3,DOSYO,DOSMO,DOSDO,
     *    DOSYS,DOSMS,DOSDS
C
          WRITE(2,200) SSAN1,SSAN2,SSAN3,DOSYO,DOSMO,DOSDO,
     *    DOSYS,DOSMS,DOSDS
 300   CONTINUE
       STOP
 100   FORMAT(3A3,164X,I2,I2,I2,47X,I2,I2,I2)
 200   FORMAT(3A3,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2)
       END
/*
//GO.FT01F001 DD UNIT=3330V,VOL=SER=MS0262,DISP=(OLD),
//    DCB=(RECFM=FB,LRECL=326,BLKSIZE=12714),
//    DSNAME=MSS.S1972.CFT79
//GO.FT02F001 DD UNIT=3350,VOL=SER=MVS004,DISP=(SHR),
//    SPACE=(CYL,(4,4)),DCB=(RECFM=FB,LRECL=27,BLKSIZE=19062),
//    DSNAME=S1972.C790
//
```

```
C THIS PROGRAM STRIPS THOSE VARS THAT CHANGE OVER THE TERM OF ENLIST-
C MENT.
//DTHOMAS JOB (1972,0398),'D.THOMAS',CLASS=B
//*MAIN ORG=NPGVM1.1972P
//  EXEC FORTXCG
//FORT.SYSIN DD *
        INTEGER*4    SSAN1,SSAN2,SSAN3,
     *  TAFMO,HYECO,PGO,MSO,DEPO,ETSYO,ETSMO,CHSVO,
     *  REO,TAFMS,HYECS,PGS,MSS,DEPS,ETSYS,ETSMS,CHSVS,
     *  RES,TAFML,HYECL,PGL,MSL,DEPL,
     *  ETSYL,ETSML,CHSVL,REL
        INTEGER*4 BLANK,NINE
        DATA BLANK/'    '/,NINE/' -9'/
      DO 300 I=1,30778,10
C
      SSAN1    = BLANK
      SSAN2    = BLANK
      SSAN3    = NINE
      TAFMO    = -9
      HYECO    = -9
      PGO      = -9
      MSO      = -9
      DEPO     = -9
      ETSYO    = -9
      ETSMO    = -9
      CHSVO    = -9
      REO      = -9
      TAFMS    = -9
      HYECS    = -9
      PGS      = -9
      MSS      = -9
      DEPS     = -9
      ETSYS    = -9
      ETSMS    = -9
      CHSVS    = -9
      RES      = -9
      TAFML    = -9
      HYECL    = -9
      PGL      = -9
      MSL      = -9
      DEPL     = -9
      ETSYL    = -9
      ETSML    = -9
      CHSVL    = -9
      REL      = -9
C
        READ(1,100) SSAN1,SSAN2,SSAN3,TAFMO,HYECO,PGO,MSO,
     *  DEPO,ETSYO,ETSMO,CHSVO,REO,
     *  TAFMS,HYECS,PGS,MSS,DEPS,ETSYS,ETSMS,CHSVS,RES,
     *  TAFML,HYECL,PGL,MSL,DEPL,ETSYL,ETSML,CHSVL,REL
C
        WRITE(2,200)SSAN1,SSAN2,SSAN3,TAFMO,HYECO,PGO,MSO,
     *  DEPO,ETSYO,ETSMO,CHSVO,REO,
     *  TAFMS,HYECS,PGS,MSS,DEPS,ETSYS,ETSMS,CHSVS,RES,
     *  TAFML,HYECL,PGL,MSL,DEPL,ETSYL,ETSML,CHSVL,REL
 300  CONTINUE
      STOP
 100  FORMAT(3A3,142X,I3,6X,I2,I2,1X,I1,I2,17X,I2,I2,4X,I1,I1,
     *11X,I3,6X,I2,I2,1X,I1,I2,17X,I2,I2,4X,I1,I1,11X,I3,
     *6X,I2,I2,1X,I1,I2,17X,I2,I2,4X,I1,I1)
 200  FORMAT(3A3,1X,
     *I3,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I3,1X,I2,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I3,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,
     *I2)
      END
/*
//GO.FT01F001 DD UNIT=3330V,VOL=SER=MS0282,DISP=(OLD),
//   DCB=(RECFM=FB,LRECL=326,BLKSIZE=12714),
//   DSNAME=MSS.S1972.CHT79
//GO.FT02F001 DD UNIT=3350,VOL=SER=MVS004,DISP=(SHR),

//   SPACE=(CYL,(4,4)),DCB=(RECFM=FB,LRECL=93,BLKSIZE=19065),
//   DSNAME=S1972.C79A
//
```

```
C THIS PROGRAM STRIPS PERSONAL VAR FROM ENTIRE 79 COHORT.
//DTHOMAS JOB (1972,0398),'D.THOMAS',CLASS=B
//*MAIN ORG=NPGVM1.1972P
//   EXEC FORTXCG
//FORT.SYSIN DD *
         INTEGER*4    SSAN1,SSAN2,SSAN3,AGEE,HYEC,SEX,RACE,ETH,
     *   RETH,MS,AFQTTG,ENTST,HT,WT,WAIV,WAIVAL,EXAMST,
     *   TOE,EPG,BONUS,TRMOS3
         INTEGER*4 BLANK,NINE
         DATA BLANK/'   '/,NINE/' -9'/
      DO 300 I=1,30778,10
C
      SSAN1    =  BLANK
      SSAN2    =  BLANK
      SSAN3    =  NINE
      AGEE     =  -9
      HYEC     =  -9
      SEX      =  -9
      RACE     =  -9
      ETH      =  -9
      RETH     =  -9
      MS       =  -9
      AFQTTG   =  -9
      ENTST    =  -9
      HT       =  -9
      WT       =  -9
      WAIV     =  -9
      WAIVAL   =  -9
      EXAMST   =  -9
      TOE      =  -9
      EPG      =  -9
      BONUS    =  -9
      TRMOS3   =  -9
C
         READ(1,100)SSAN1,SSAN2,SSAN3,AGEE,HYEC,SEX,RACE,ETH,
     *   RETH,MS,AFQTTG,ENTST,HT,WT,WAIV,WAIVAL,EXAMST,
     *   TOE,EPG,BONUS,TRMOS3
C
         WRITE(2,200)SSAN1,SSAN2,SSAN3,AGEE,HYEC,SEX,RACE,ETH,
     *   RETH,MS,AFQTTG,ENTST,HT,WT,WAIV,WAIVAL,EXAMST,
     *   TOE,EPG,BONUS,TRMOS3
 300  CONTINUE
      STOP
 100  FORMAT(3A3,20X,I2,1X,I2,I1,I1,I2,I1,I2,4X,I1,40X,
     *I1,I2,I3,12X,I2,
     *I2,I2,6X,I2,I2,13X,I1,13X,I4)
 200  FORMAT(3A3,1X,
     *I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I2,1X,I3,1X,I2,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I2,1X,I4)
      END
/*
//GO.FT01F001 DD UNIT=3330V,VOL=SER=MSS282,DISP=(OLD),
//    DCB=(RECFM=FB,LRECL=326,BLKSIZE=12714),
//    DSNAME=MSS.S1972.CHT79
//GO.FT02F001 DD UNIT=3350,VOL=SER=MVS004,DISP=(NEW,KEEP),
//    SPACE=(CYL,(4,4)),DCB=(RECFM=FB,LRECL=66,BLKSIZE=19008),
//    DSNAME=S1972.C79B
//
```

```
C
C
//DTHOMAS JOB (1972,0398),'D.THOMAS',CLASS=B
//*MAIN ORG=NPGVM1.1972P
//   EXEC FORTXCG
//FORT.SYSIN DD *
      INTEGER*4     SSAN1,SSAN2,SSAN3,
     *  CENR,CEND,ZIP1,ZIP2,HOR,RECID,TF,AFQTP,APTA,
     *  APTB,APTC,APTD,APTE,APTF,APTG,APTH,APTI,APTJ,APTK,APTL,
     *  APTM,APTN,APTO,APTP,AFEES
      INTEGER*4 BLANK,NINE
      DATA BLANK/'    '/,NINE/' -9'/,NIN/'-9'/
     DO 300 I=1,30778,10
C
     SSAN1    = BLANK
     SSAN2    = BLANK
     SSAN3    = NINE
     CENR     = -9
     CEND     = -9
     ZIP1     = BLANK
     ZIP2     = NIN
     HOR      = -9
     RECID    = -9
     TF       = -9
     AFQTP    = -9
     APTA     = -9
     APTB     = -9
     APTC     = -9
     APTD     = -9
     APTE     = -9
     APTF     = -9
     APTG     = -9
     APTH     = -9
     APTI     = -9
     APTJ     = -9
     APTK     = -9
     APTL     = -9
     APTM     = -9
     APTN     = -9
     APTO     = -9
     APTP     = -9
     AFEES    = -9
C
      READ(1,100) SSAN1,SSAN2,SSAN3,
     *  CENR,CEND,ZIP1,ZIP2,HOR,
     *  RECID,TF,AFQTP,APTA,APTB,
     *  APTC,APTD,APTE,APTF,
     *  APTG,APTH,APTI,APTJ,APTK,
     *  APTL,APTM,APTN,APTO,APTP
C
      WRITE(2,200) SSAN1,SSAN2,SSAN3,
     *  CENR,CEND,ZIP1,ZIP2,HOR,
     *  RECID,TF,AFQTP,APTA,APTB,
     *  APTC,APTD,APTE,APTF,
     *  APTG,APTH,APTI,APTJ,APTK,
     *  APTL,APTM,APTN,APTO,APTP,AFEES
 300  CONTINUE
      STOP
 100  FORMAT(3A3,I2,I1,A3,A2,I2,12X,I1,9X,I2,I2,1X,I2,I2,I2,I2,I2,I2,
     *I2,I2,I2,I2,I2,I2,8X,I2,I2,I2,I2,40X,I2)
 200  FORMAT(3A3,1X,
     *I2,1X,I2,1X,2A3,1X,I2,1X,I2,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,
     *I2,1X,I2,1X,I2)
      END
/*
//GO.FT01F001 DD UNIT=3330V,VOL=SER=MS0282,DISP=(OLD),
//    DCB=(RECFM=FB,LRECL=326,BLKSIZE=12714),
//    DSNAME=MSS.S1972.CFT79
//GO.FT02F001 DD UNIT=3330,VOL=SER=MVS004,DISP=(SHR),
//    SPACE=(CYL,(4,4)),DCB=(RECFM=FB,LRECL=85,BLKSIZE=19040),

//    DSNAME=S1972.C79C
//
```

APPENDIX D

APL PROGRAMS FOR DATA MANIPULATION

```
       ∇STRIP2[⎕]∇
      ∇ STRIP2
[1]    TAFMO←DRAFT1[;1]
[2]    HYECO←DRAFT1[;2]
[3]    RGO←DRAFT1[;3]
[4]    MSO←DRAFT1[;4]
[5]    DEFO←DRAFT1[;5]
[6]    CHSVO←DRAFT1[;6]
[7]    REO←DRAFT1[;7]
[8]    AGEE←DRAFT1[;8]
[9]    HYEC←DRAFT1[;9]
[10]   SEX←DRAFT1[;10]
[11]   RACE←DRAFT1[;11]
[12]   ETH←DRAFT1[;12]
[13]   MS←DRAFT1[;13]
[14]   AFOTTG←DRAFT1[;14]
[15]   MOS←DRAFT1[;15]
[16]   A←DRAFT2[;2]
[17]   B←DRAFT2[;3]
[18]   C←DRAFT3[;4]
[19]   D←DRAFT3[
      ∇
```

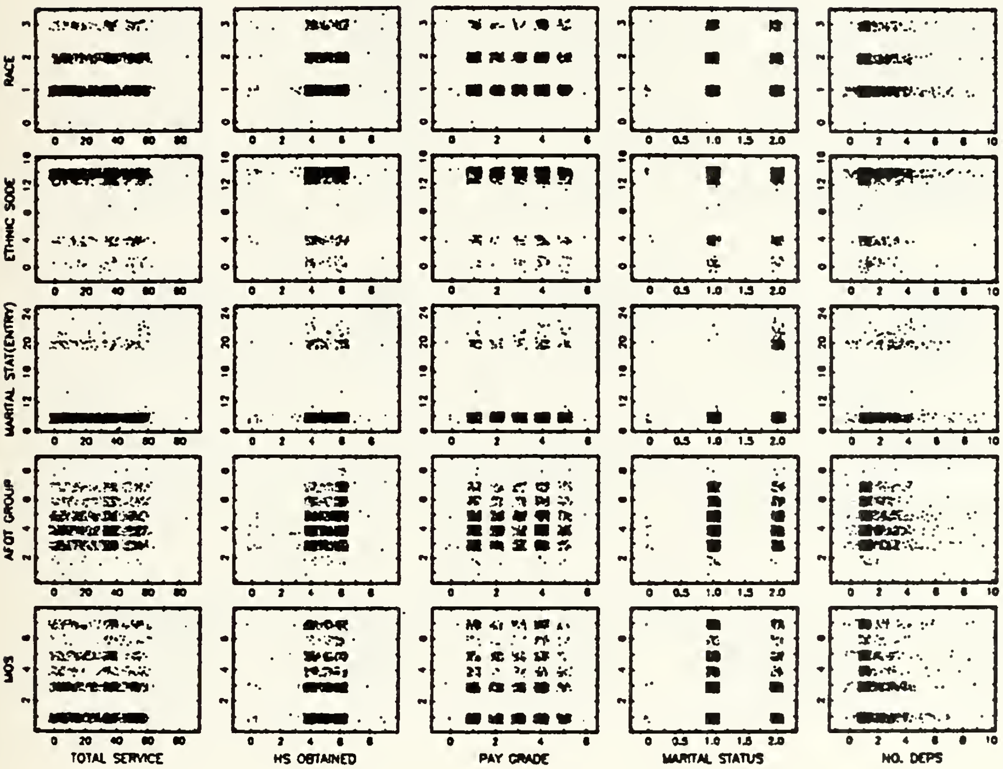
```
       ∇RECODE[⎕]∇
      ∇ RECODE DATA;C;D;NCODE;TSTOR;RDATA
[1]    ⎕←'INSERT THE COLUMN TO BE RECODED'
[2]    C←⎕
[3]    ⎕←'INSERT THE DIGITS TO BE RECODED'
[4]    D←⎕
[5]    ⎕←'INSERT THE NUMBER TO BE RECODED TO'
[6]    NCODE←⎕
[7]    TSTOR←DATA
[8]    DATA[;C]←DATA[;C]=D
[9]    DATA[;C]←DATA[;C]×NCODE
[10]   RDATA←DATA
[11]   RDATA[;C]←RDATA[;C]+TSTOR[;C]×(RDATA[;C]≠NCODE)
[12]   REDATA←RDATA
[13]   ⎕←'THE RECODED DATA IS NOW A GLOBAL VARIABLE CALLED REDATA'
      ∇
```

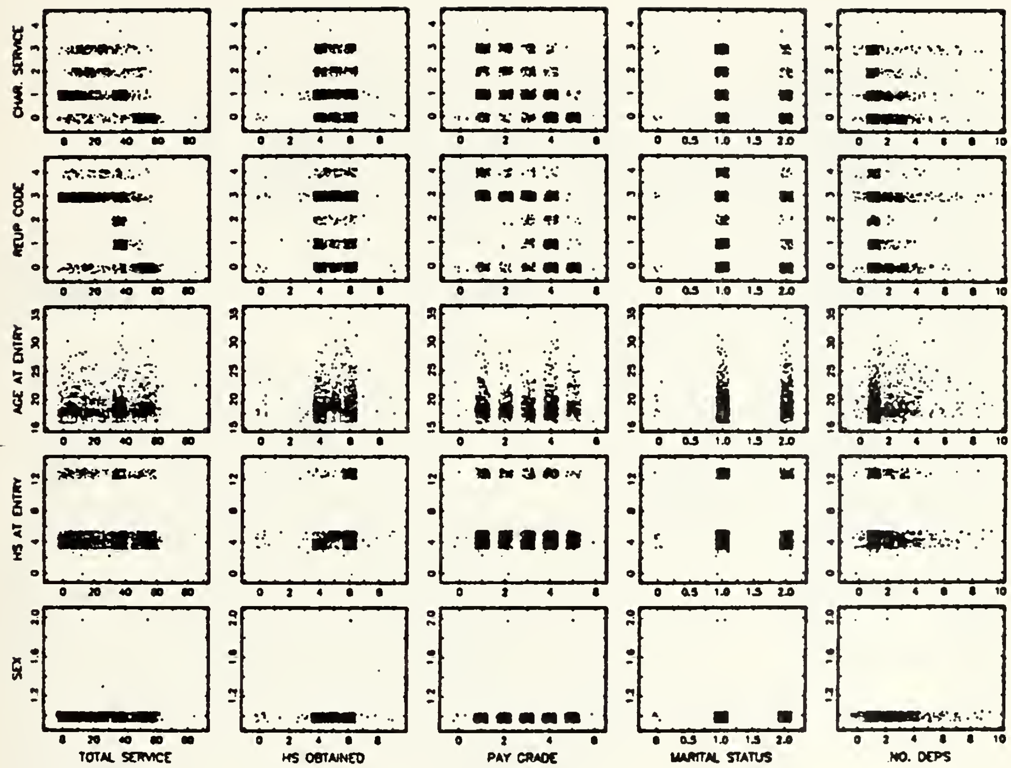
```
     ∇AVESCAN[⎕]∇
     ∇ R←AVESCAN ARRAY;C1;C2;A;VAL;V
[1]   ⍝THIS FUNCTION CHOOSES A SUBSET OF AN ARRAY AND
[2]   ⍝CALCULATES THE MEAN OF THAT SUBSET
[3]    ⎕←'INSERT THE COLUMN NUMBER OF THE SELECTION VECTOR'
[4]    C1←⎕
[5]    V←ARRAY[;C1]
[6]    ⎕←'INSERT THE COLUMN NUMBER OF OF THE VAR YOU WANT AVERAGED'
[7]    C2←⎕
[8]    A←ARRAY[;C2]
[9]    ⎕←'INSERT THE DESIRED SELECTION VALUE FOR THE SEL VECTOR'
[10]   VAL←⎕
[11]   SEL←(V=VAL)/A
[12]   ⎕←'THE SUBSET THAT YOU HAVE SELECTED IS STORED AS GLOBAL VARIABLE
       SEL'
[13]   ⎕←'THE AVERAGE OF YOUR SELECTED PERSONNEL IS'
[14]   R←(+/SEL)÷(⍴SEL)
     ∇

     ∇COLAT[⎕]∇
     ∇ COLAT
[1]    DRAFT1←A79[;2],A79[;3],A79[;4],A79[;5],A79[;6],A79[;9],A79[;10]
[2]    DRAFT1←DRAFT1,B79[;2],B79[;3],B79[;4],B79[;5],B79[;6],B79[;8]
[3]    DRAFT1←DRAFT1,B79[;9],B79[;19]
[4]    DRAFT2←A79[;2],C79[;9],C79[;10],C79[;11],C79[;12],C79[;13],C79[;14
       ]
[5]    DRAFT2←DRAFT2,C79[;15],C79[;16],C79[;17],C79[;18],C79[;19],C79[;20
       ]
[6]    DRAFT2←DRAFT2,C79[;21],C79[;22],C79[;23],C79[;24]
[7]    DRAFT1←⍉(15 3078 ⍴DRAFT1)
[8]    DRAFT2←⍉(17 3078 ⍴DRAFT2)
     ∇

     ∇MATBLD[⎕]∇
     ∇ MATBLD
[1]   ⍝BUILDS MATRIX FOR USE IN ANOVA TESTING IN CHAPTER 3
[2]    ROW← ¯99999 ¯99999 ¯99999 ¯99999 ¯99999
[3]    MATR← 1543 5 ⍴ROW
[4]    MATR[;1]←GED,(196↓MATR[;1])
[5]    MATR[;2]←HSDG
[6]    MATR[;3]←NHSG,(503↓MATR[;3])
[7]    MATR[;4]←JR,(751↓MATR[;4])
[8]    MATR[;5]←SOPH,(10↓MATR[;5])
     ∇
```

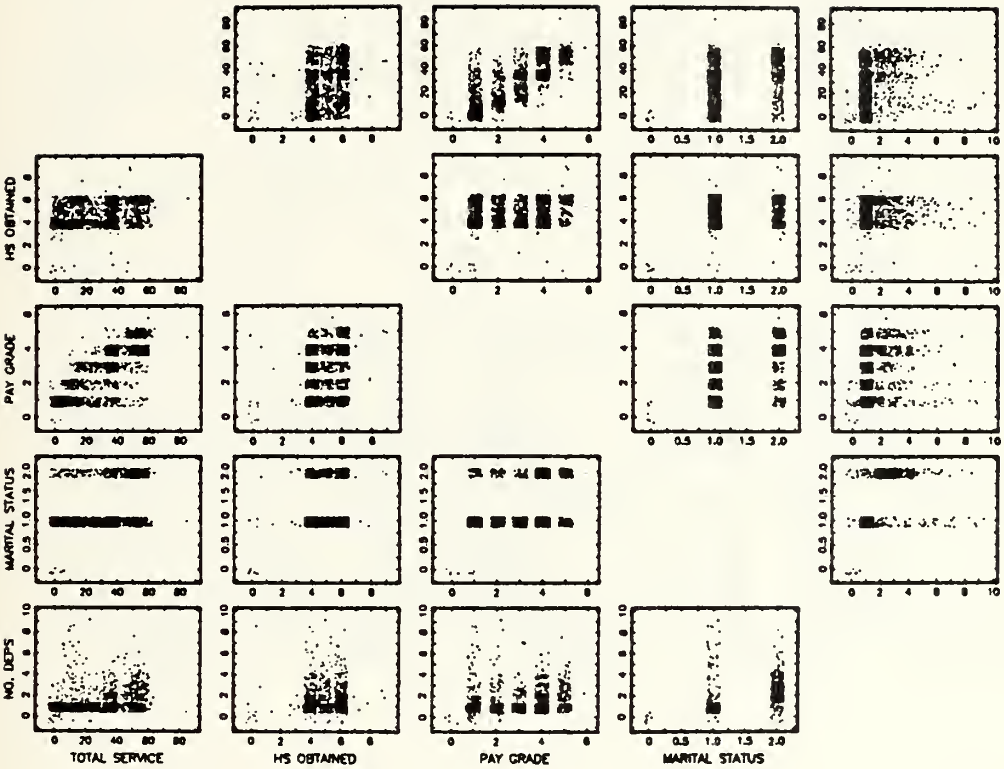
```
      ∇PRC[⎕]∇
     ∇ P←PRC DATA;DATA1;SEL;NEQ36;NLT36;NGT36;EQ36;LT36;GT36;PEQ36;PLT3
       6;PGT36
[1]   ⍝CALCULATES REL.FREQ(PROBABILITY) FOR SURVIVAL ANALYSIS.
[2]    ⎕←'CHECK TO SEE IF THIS FN IS SET UP FOR 36 OF 48 YO'
[3]    AVESCAN DATA
[4]    DATA1←SEL
[5]    NEQ48←ρEQ48←(DATA1=48)/DATA1
[6]    NLT48←ρLT48←(DATA1<48)/DATA1
[7]    NGT48←ρGT36←(DATA1>48)/DATA1
[8]    P2←PEQ48←NEQ48÷ρSEL
[9]    P1←PLT48←NLT48÷ρSEL
[10]   P3←PGT48←NGT48-ρSEL
[11]   ⎕←'VECTOR OF PROBABILITIES:P(LOS<48),P(LOS=48),P(LOS>48)'
[12]   P←P1,P2,P3
     ∇


      ∇YO4MKR[⎕]∇
     ∇ YO4MKR DATA
[1]   ⍝BUILDS MATRIX OF 4YO ONLY DATA, DATA IS ORIGINAL
[2]   ⍝DATA SET.
[3]    YO4←DATA[;11]=4
[4]    NDATA4← 479 10 ρ(1 1 1 1 1 1 1 1 1 1)
[5]    NDATA4[;1]←YO4/DATA[;1]
[6]    NDATA4[;2]←YO4/DATA[;2]
[7]    NDATA4[;3]←YO4/DATA[;3]
[8]    NDATA4[;4]←YO4/DATA[;4]
[9]    NDATA4[;5]←YO4/DATA[;5]
[10]   NDATA4[;6]←YO4/DATA[;6]
[11]   NDATA4[;7]←YO4/DATA[;7]
[12]   NDATA4[;8]←YO4/DATA[;8]
[13]   NDATA4[;9]←YO4/DATA[;9]
[14]  .NDATA4[;10]←YO4/DATA[;10]
     ∇
      ∇YO3MKR[⎕]∇
     ∇ YO3MKR DATA
[1]   ⍝BUILDS MATRIX OF 3YO ONLY DATA, 'DATA' IS ORIGINAL
[2]   ⍝DATA SET
[3]    YO3←DATA[;11]=3
[4]    NDATA← 2565 10 ρ(1 1 1 1 1 1 1 1 1 1)
[5]    NDATA[;1]←YO3/DATA[;1]
[6]    NDATA[;2]←YO3/DATA[;2]
[7]    NDATA[;3]←YO3/DATA[;3]
[8]    NDATA[;4]←YO3/DATA[;4]
[9]    NDATA[;5]←YO3/DATA[;5]
[10]   NDATA[;6]←YO3/DATA[;6]
[11]   NDATA[;7]←YO3/DATA[;7]
[12]   NDATA[;8]←YO3/DATA[;8]
[13]   NDATA[;9]←YO3/DATA[;9]
[14]   NDATA[;10]←YO3/DATA[;10]
     ∇
```

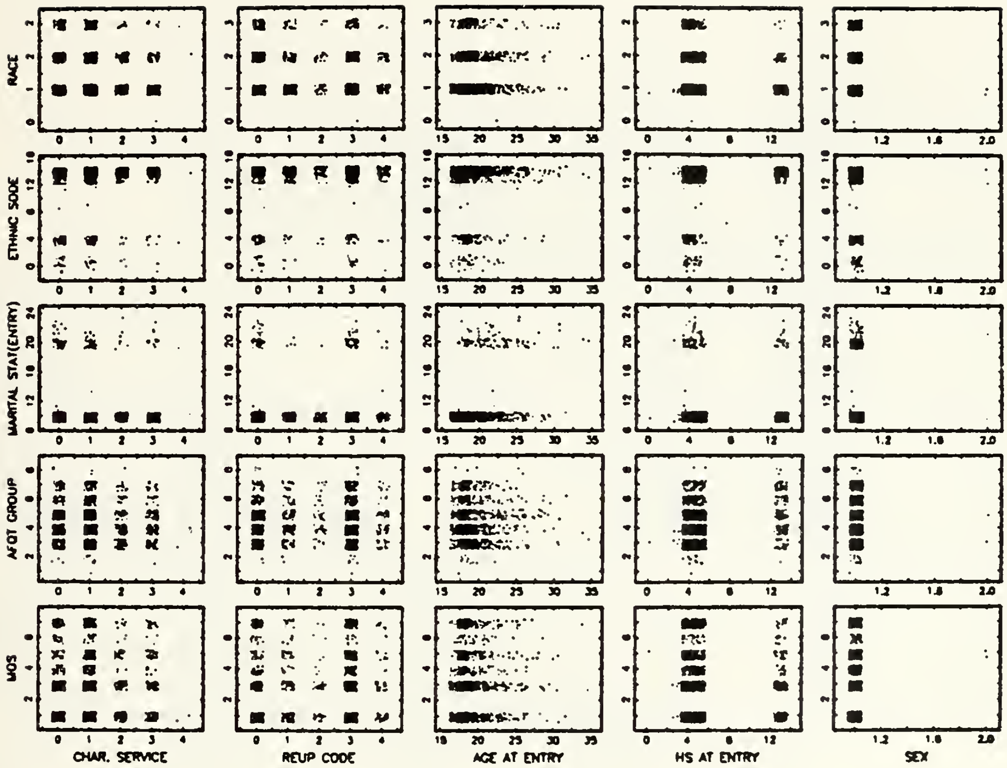
APPENDIX E

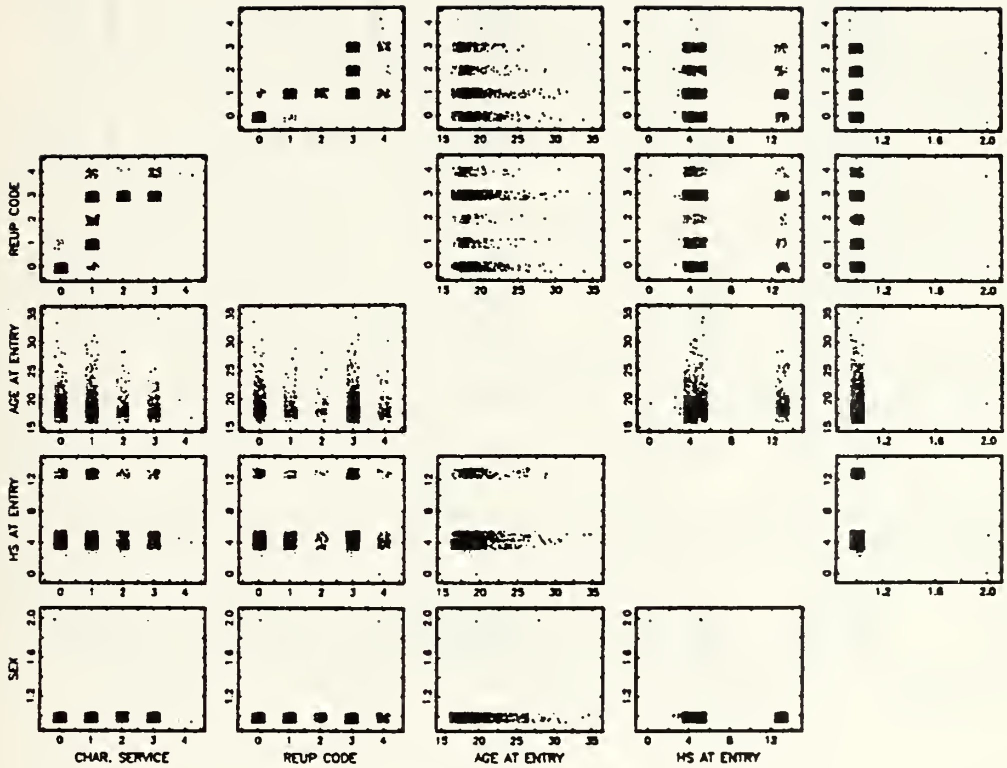
OVERALL VIEW OF FIRST DRAFTSMAN'S DISPLAY

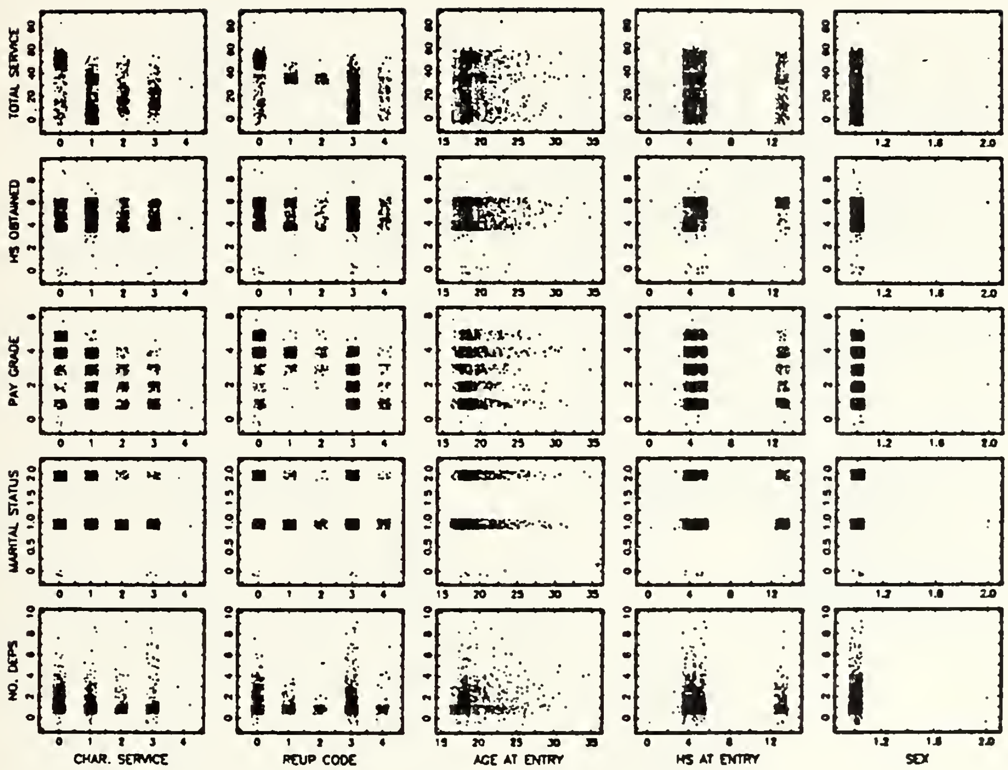


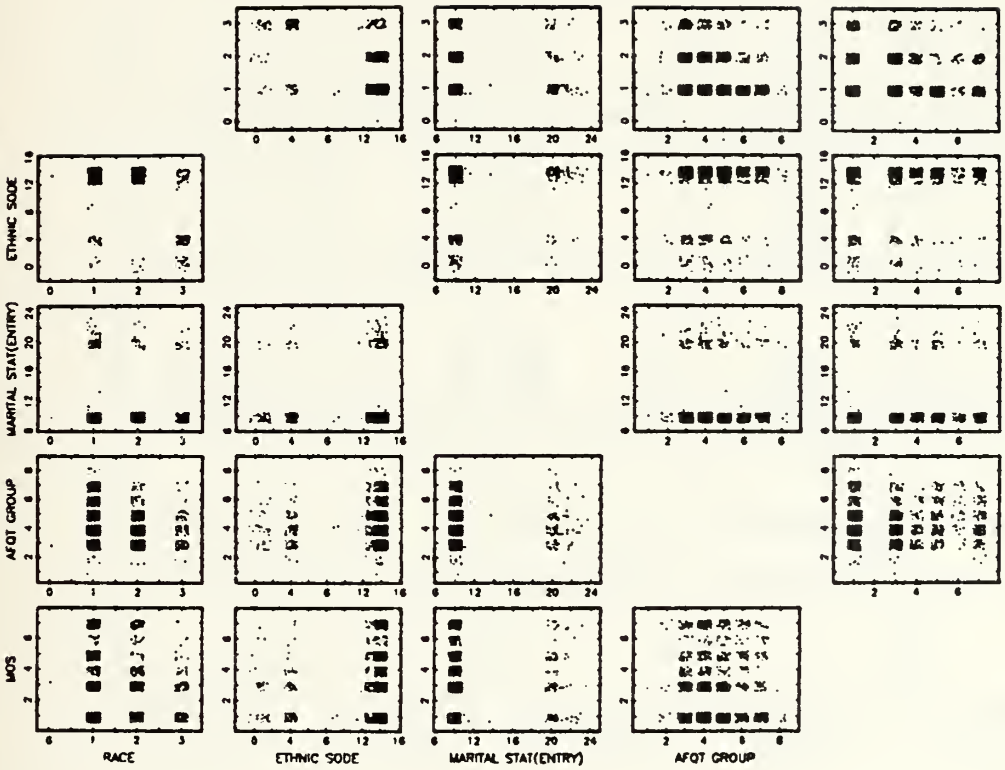

CHAR. SERVICE
REUP CODE
AGE AT ENTRY
HS AT ENTRY
SEX
TOTAL SERVICE
HS OBTAINED
PAY GRADE
MARITAL STATUS
NO. DEPS


HS OBTAINED
PAY GRADE
MARITAL STATUS
NO. DEPS
TOTAL SERVICE
HS OBTAINED
PAY GRADE
MARITAL STATUS

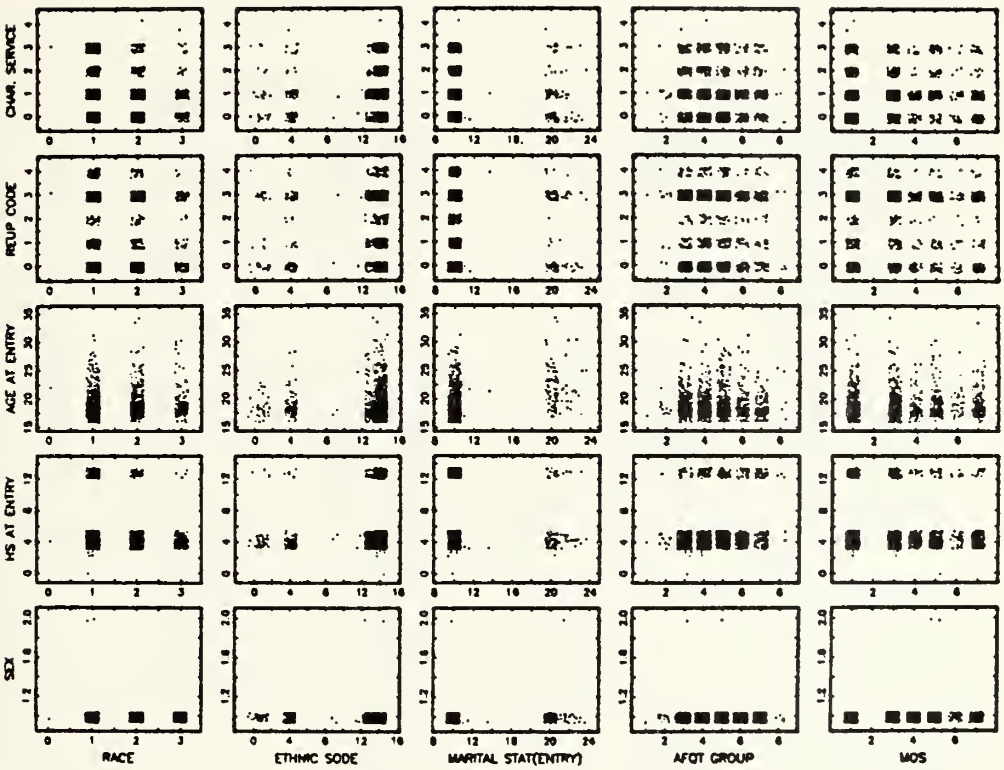

RACE
ETHNIC SODE
MARITAL STAT(ENTRY)
AFQT GROUP
MOS
CHAR. SERVICE
REUP CODE
AGE AT ENTRY
HS AT ENTRY
SEX

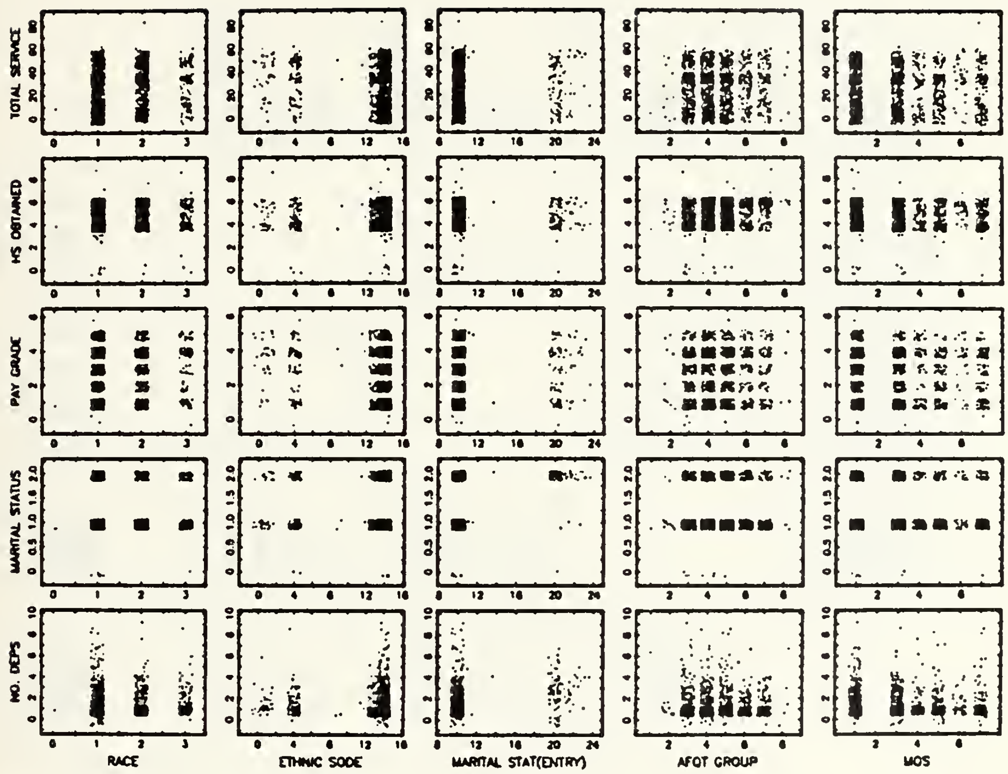

REUP CODE
AGE AT ENTRY
HS AT ENTRY
SEX
CHAR. SERVICE
REUP CODE
AGE AT ENTRY
HS AT ENTRY


TOTAL SERVICE
HS OBTAINED
PAY GRADE
MARITAL STATUS
NO. DEPS
CHAR. SERVICE
REUP CODE
AGE AT ENTRY
HS AT ENTRY
SEX



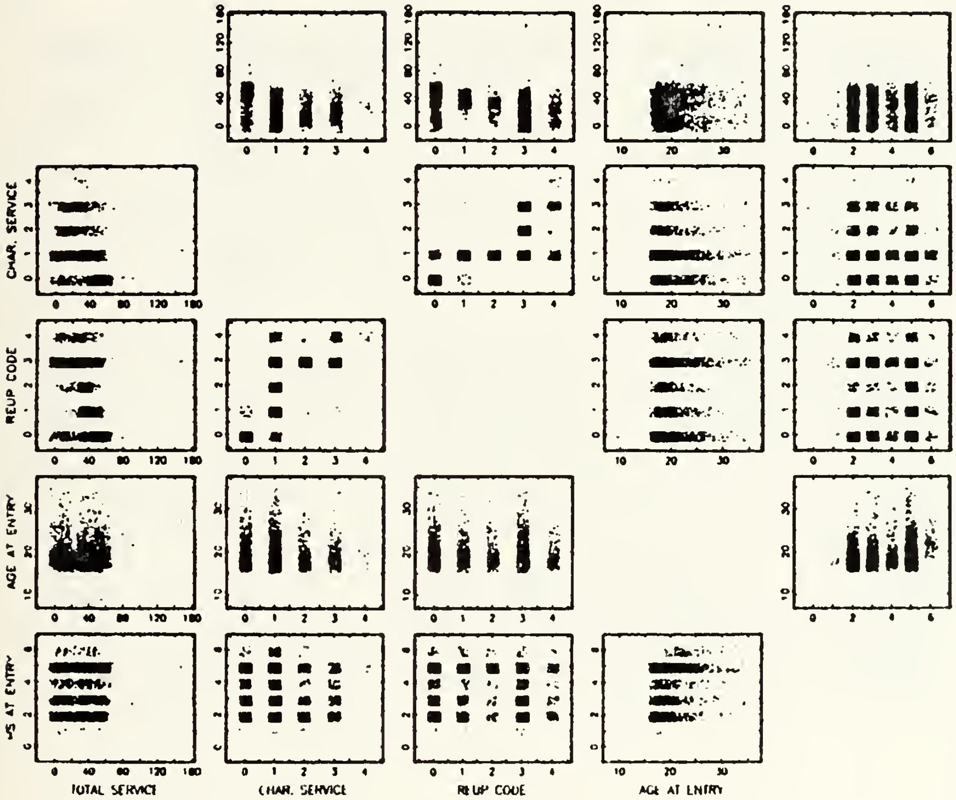
ETHNIC SODE
MARITAL STAT(ENTRY)
AFQT GROUP
MOS
RACE
ETHNIC SODE
MARITAL STAT(ENTRY)
AFQT GROUP

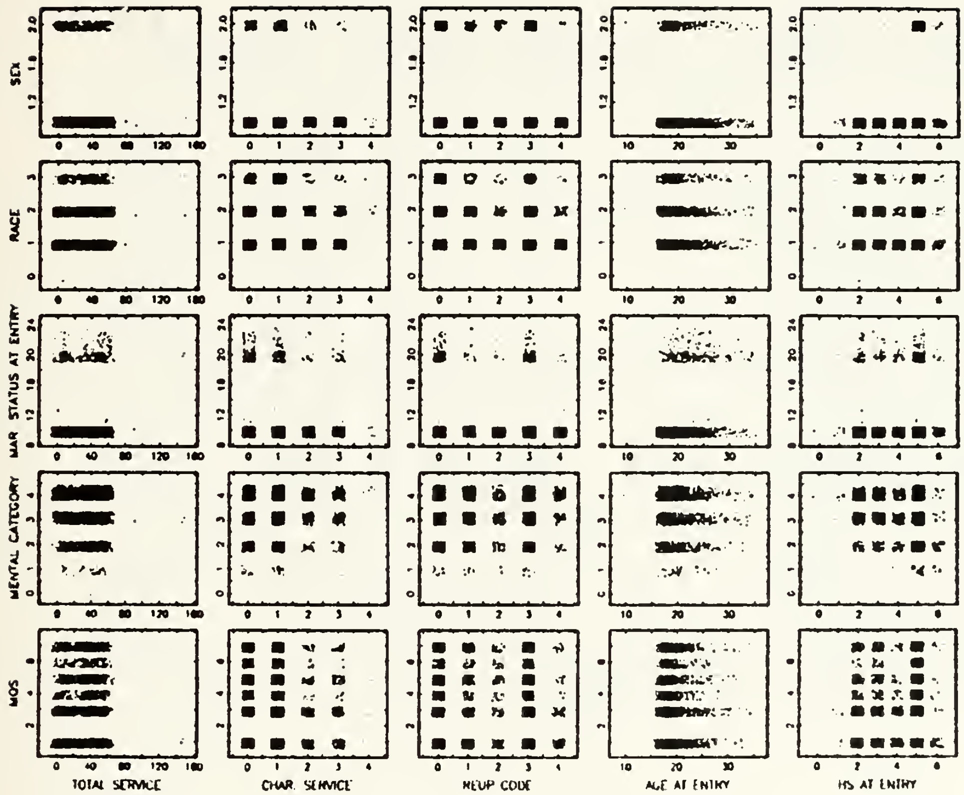

CHAR. SERVICE
REUP CODE
AGE AT ENTRY
HS AT ENTRY
SEX
RACE
ETHNIC SODE
MARITAL STAT(ENTRY)
AFQT GROUP
MOS

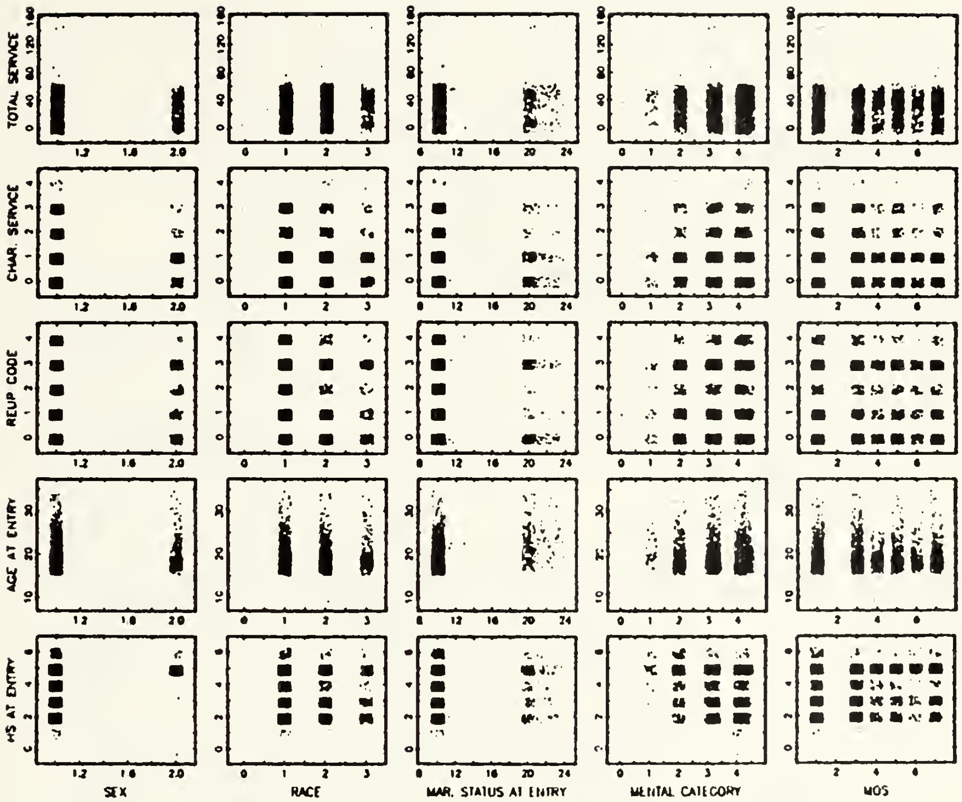

TOTAL SERVICE
HS OBTAINED
PAY GRADE
MARITAL STATUS
NO. DEPS
RACE
ETHNIC SODE
MARITAL STAT(ENTRY)
AFQT GROUP
MOS

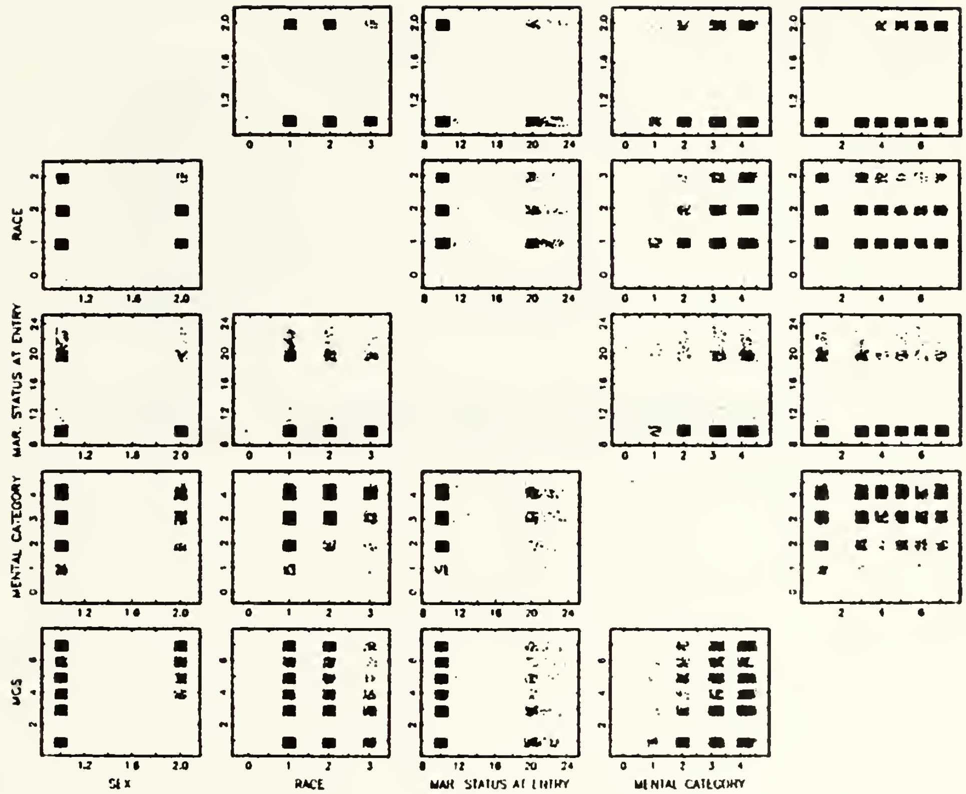
APPENDIX F

OVERALL VIEW OF REVISED DRAFTSMAN'S DISPLAY




SEX
RACE
MAR STATUS AT ENTRY
MENTAL CATEGORY
MOS
TOTAL SERVICE
CHAR. SERVICE
REUP CODE
AGE AT ENTRY
HS AT ENTRY


TOTAL SERVICE
CHAR. SERVICE
REUP CODE
AGE AT ENTRY
HS AT ENTRY
SEX
RACE
MAR. STATUS AT ENTRY
MENTAL CATEGORY
MOS


RACE
MAR. STATUS AT ENTRY
MENTAL CATEGORY
MCS
SEX
RACE
MAR. STATUS AT ENTRY
MENTAL CATEGORY

APPENDIX G

INPUT SCREEN, IBM GRAFSTAT SUBPOPULATION CATEGORY ANALYSIS PROGRAM

```
                    CATEGORY  ANALYSIS

CATEGORY  VECTOR               : C
Y  VARIABLE                    : Y
UNWEIGHTED (0) OR WEIGHTS      : 0
SELECTION                      : Δ

PERCENTILES (OR BOX)           : BOX
SYMBOLS DEFINED (Y/N)          : N
PLOT HEADER (IN QUOTES)        : Δ

SCREEN HEADER (IN QUOTES)      : ''
X-AXIS LABEL (IN QUOTES)       : Δ
Y-AXIS LABEL (IN QUOTES)       : Δ
PLOT POSITION    :                          TABLE POSITION:
SCALE X-AXIS     LIN                        SCALE Y-AXIS  : LIN
PARTIAL PLOT                     1 1 1
AXES AND GRID CONTROL            0 1 0 0


ENTER=GO          PF:1=HELP  2=VIEW GRAPHICS(3279)  3=RETURN  4=WRITE ON SCREEN
CLEAR=DEFAULT        5=LAST RESPONSES     6=ERASE      7=PROFILE     8=UNDEFINED
     RESPONSES       9=OUTPUT  10=STORE/RETRIEVE  11=INTO APL 12=SCREEN DISPLAY
```

APPENDIX H

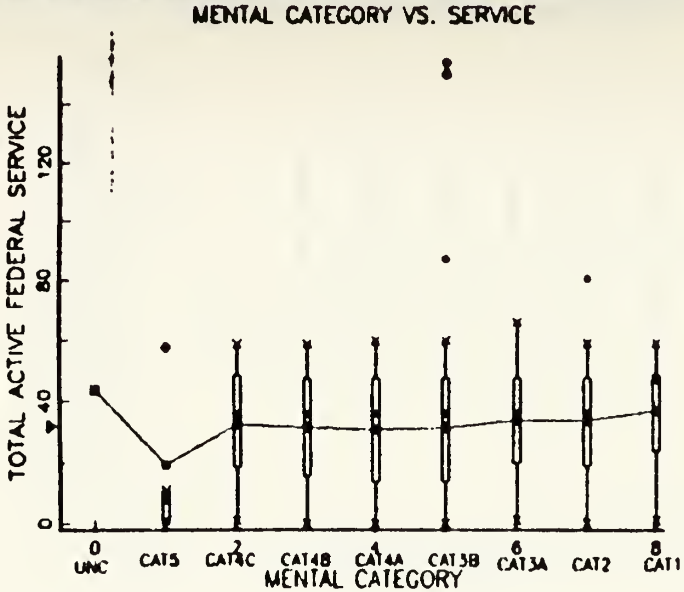
BOXPLOT ANALYSIS OF REMAINING VARIABLES

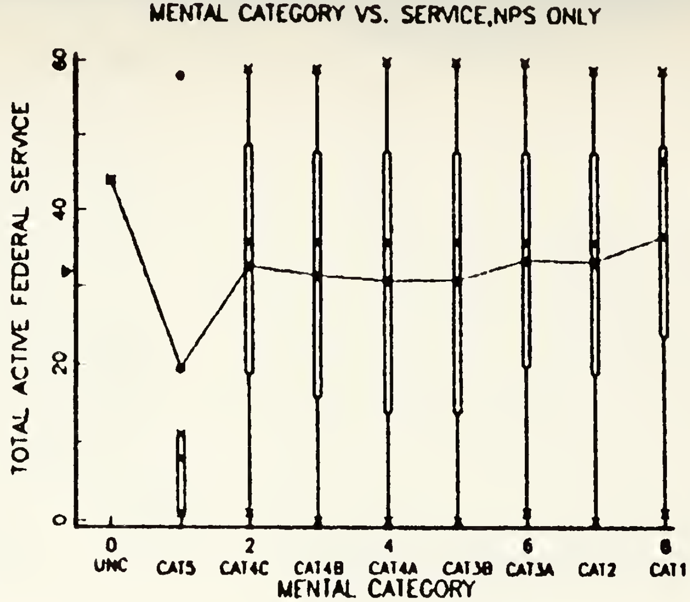
Boxplots of the remaining candidate explanatory variables versus length of service are provided in this appendix. Refer to Chapter 3, pages 62 through 66 for discussion on each of these boxplots. Remaining candidate explanatory variables displayed in this appendix are listed in Table XVI

TABLE XVI

Candidate Explanatory Variables

Mental Category
Marital Status
Age
Sex
Race
Reenlistment Code
Character of Service





CATEGORY VECTOR :MENCT
Y :Y-A79MOD[;1]
SELECTION :ALL

X LABEL :MENT
Y LABEL :TOTA
NO. OF ELEMENTS :3078

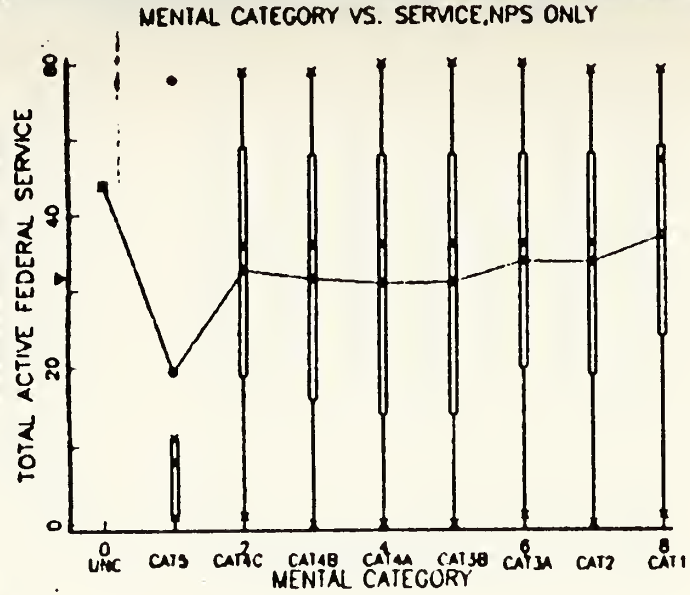
CATEG.	NO PTS	I*/*-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	31.875	17.836	16	36	48
0	1	0.00032489	44	0	44	44	44
1	4	0.0012995	19.5	22.522	1	8	11
2	230	0.074724	32.704	17.778	19	36	49
3	613	0.19916	31.515	17.35	16	36	48
4	796	0.25861	30.898	18.021	14	36	48
5	681	0.22125	31.423	18.798	14	36	46
6	330	0.10721	33.703	16.088	20	36	46
7	385	0.12508	33.605	16.903	19	36	48
8	38	0.012346	36.921	17.666	24	47	49

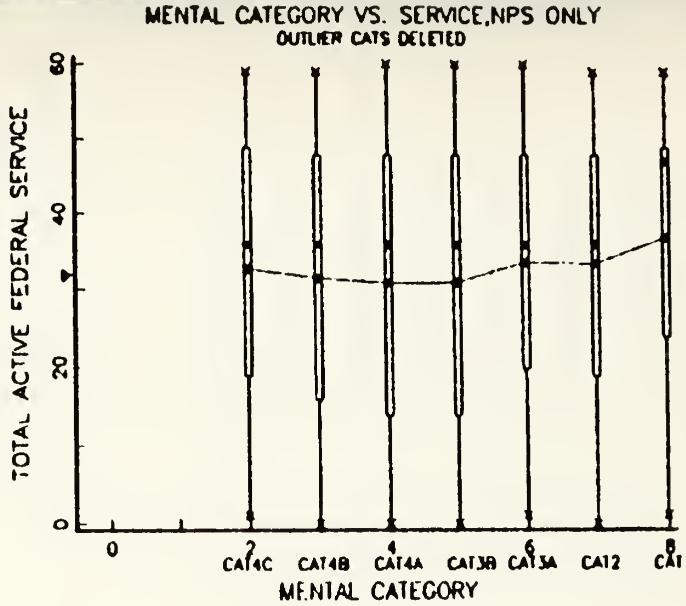
CATEGORY VECTOR :MENCT
Y :Y-A79MOD[;1]
SELECTION :Y≤60

X LABEL :MENT
Y LABEL :TOTA
NO. OF ELEMENTS :3073

CATEG.	NO.PTS	I*/*-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	36	48
0	1	0.00032541	44	0	44	44	44
1	4	0.0013017	19.5	22.522	1	8	11
2	230	0.074845	32.704	17.776	19	36	49
3	613	0.19948	31.515	17.35	16	36	48
4	798	0.25903	30.898	18.021	14	36	48
5	678	0.22063	30.985	17.531	14	36	48
6	328	0.10706	33.605	16.017	20	36	48
7	384	0.12496	33.484	16.759	18	36	48
8	38	0.012346	36.821	17.666	24	47	48

Figure H.1   Mental Category vs. Length of Service,I





CATEGORY VECTOR :MENCT
Y :Y←A79MOD[;1]
SELECTION :Y≤60

X LABEL :MENT
Y LABEL :TOTA
NO. OF ELEMENTS :3073

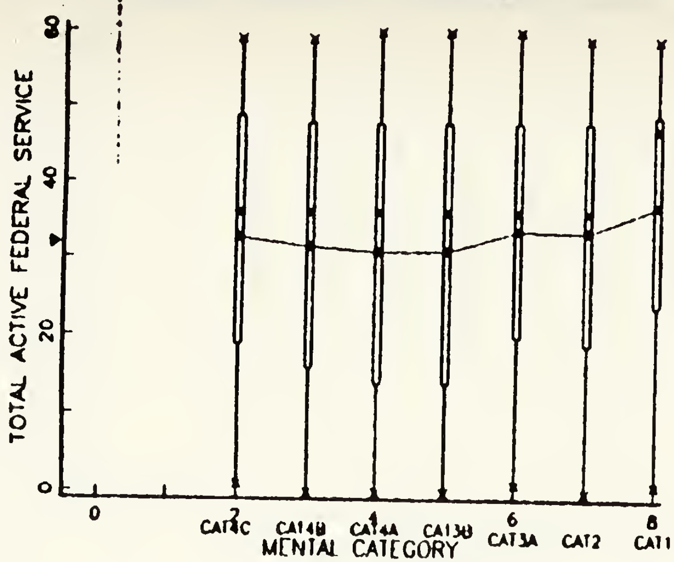
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	16	36	48
0	1	0.00032541	44	0	44	44	44
1	4	0.0013017	19.5	22.522	1	8	11
2	230	0.074845	32.704	17.776	18	36	49
3	613	0.19948	31.515	17.35	16	36	48
4	796	0.25903	30.898	18.021	14	36	48
5	678	0.22063	30.985	17.531	14	36	48
6	329	0.10706	33.805	16.817	20	36	48
7	384	0.12496	33.484	16.759	18	36	48
8	38	0.012366	36.921	17.866	24	47	49

CATEGORY VECTOR :MENCT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;6]≠0)∧(B79MOD[;6]≠1)

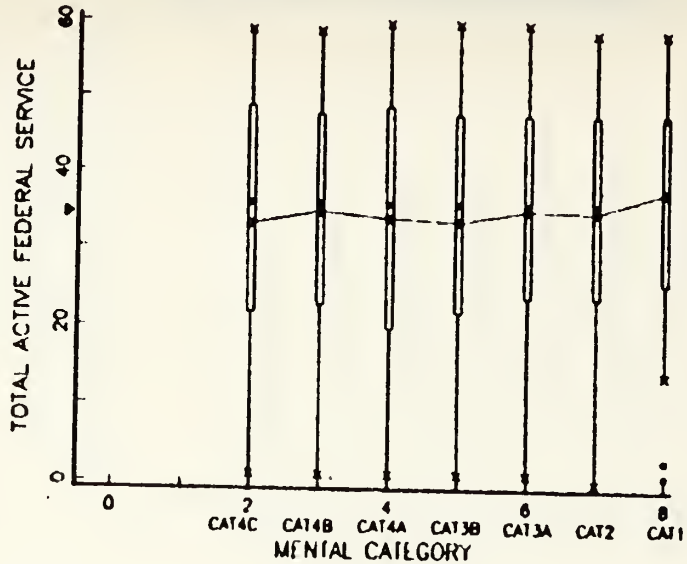
CATEG	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3068	1	31.884	17.516	16	36	48
2	230	0.074967	32.704	17.776	18	36	48
3	613	0.1998	31.515	17.35	16	36	48
4	796	0.25945	30.898	18.021	14	36	48
5	678	0.22099	30.985	17.531	14	36	48
6	329	0.10724	33.605	16.817	20	36	48
7	384	0.12516	33.484	16.759	19	36	48
8	38	0.012386	38.921	17.666	24	47	49

Figure H.2 Mental Category vs. Length of Service,II

MENTAL CATEGORY VS. SERVICE,NPS ONLY



MENTAL CATEGORY VS. SERVICE,HSDG AND HIGHER



CATEGORY VECTOR :MENCT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;6]=0)∧(B79MOD[;6]≠1)

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	75
ALL	3088	1	31.884	17.516	18	36	48
2	230	0.074967	32 704	17 776	19	36	49
3	613	0.1998	31 515	17 35	16	36	48
4	798	0 25945	30 888	18 021	14	36	48
5	678	0.22098	30 985	17 531	14	36	48
6	329	0.10724	33 605	18 817	20	36	48
7	384	0 12516	33 484	18.759	19	36	48
8	38	0 012384	36 921	17 866	24	47	49

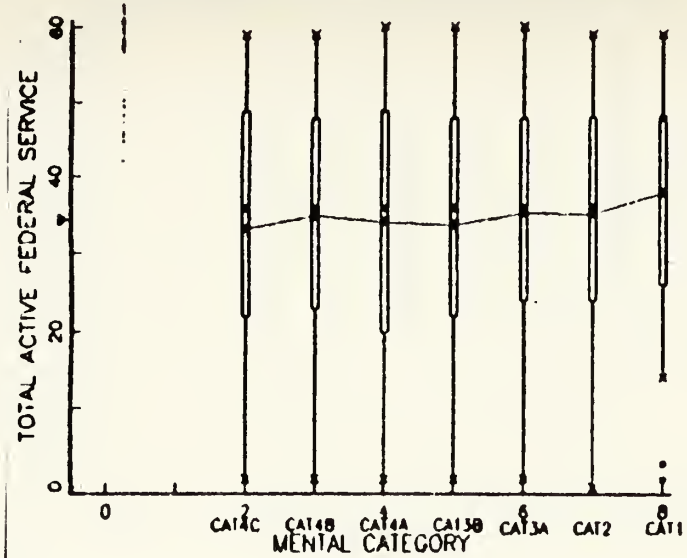
CATEGORY VECTOR :MENCT
Y :Y←A79MOD[;1]
SELECTION :MCS

X LABEL :ME
Y LABEL :TO
NO. OF ELEMENTS :10

CATEG	NO PTS	o/o-PTS	Y-MEAN	Y DVN	.25	.50	.75
ALL	1608	1	34 382	18 047	23	38	48
2	218	0 13557	33.178	17 505	22	38	49
3	267	0 16604	34 885	18 489	23	38	48
4	318	0 19776	34.05	17 802	20	38	49
5	317	0 19714	33 834	18 905	22	38	48
6	186	0 11567	35 247	18 031	24	38	48
7	269	0 16729.	35.093	18 104	24	36	48
8	33	0 020522	37 879	18 551	28	48	48

Figure H.3 Mental Category vs. Length of Service,III

MENTAL CATEGORY VS. SERVICE,HSDG AND HIGHER

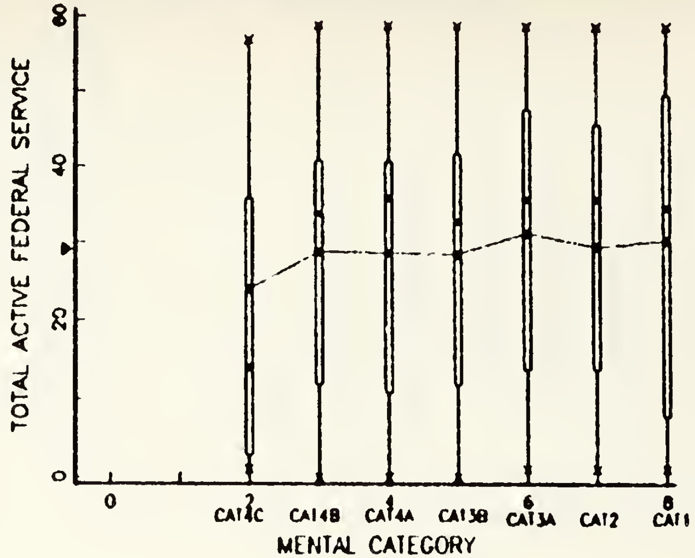


CATEGORY VECTOR :MENCT
Y :Y+A79MOD[;1]
SELECTION :MCS

X LABEL :ME
Y LABEL :TO
NO. OF ELEMENTS :16

CATEG	NO.PTS	o/o PTS	Y MEAN	Y-DVN	.25	.50	.75
ALL	1608	1	34 382	16.847	23	36	48
2	218	0 13557	33 179	17 505	22	36	49
3	267	0 16604	34 895	16 480	23	38	48
4	318	0.19776	34 05	17 802	20	38	49
5	317	0 19714	33 834	16 905	22	38	48
6	186	0 11567	35 247	18 031	24	38	48
7	269	0.16729	35 083	18 104	24	38	48
8	33	0.020522	37.879	16.551	28	48	48

MENTAL CATEGORY VS. SERVICE,NHSG AND GED



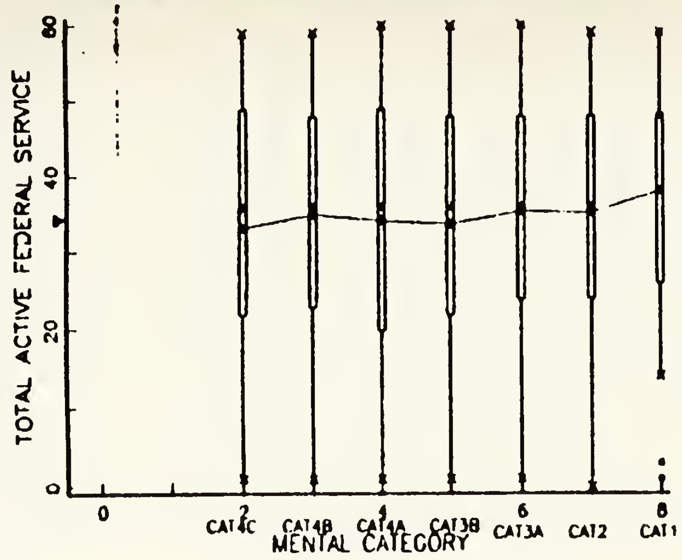
CATEGORY VECTOR :MENCT
Y :Y+A79MOD[;1]
SELECTION :MCSNC

X LABEL :MEN
Y LABEL :TOT
NO. OF ELEMENTS :148

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1460	1	29 091	17 817	12	34	42
2	12	0 0082192	24 083	20 271	3	14	36
3	346	0 23699	26 908	17 548	12	34	41
4	478	0 3274	28 797	17 99	11	34	41
5	361	0 24726	28 659	17.74	12	33	42
6	143	0 097945	31 489	17 56	14	36	48
7	115	0 078767	29.722	17 634	14	38	46
8	5	0.0034247	30.6	22.756	8	35	50

Figure H.4 Mental Category vs. Length of Service,IV

MENTAL CATEGORY VS. SERVICE,HSDG AND HIGHER

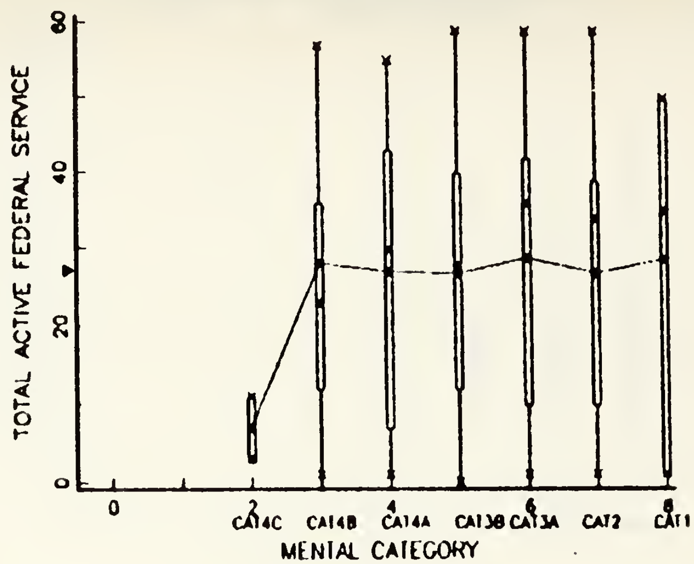


CATEGORY VECTOR :MENC1
Y :Y←A79MOD[;1]
SELECTION :MCS

X LABEL :ME
Y LABEL :TO
NO OF ELEMENTS :16

CATEG	NO.PTS	o/o-PTS	Y-MEAN	Y DVN	.25	.50	.75
ALL	1600	1	34 382	16.847	23	36	48
2	218	0 13557	33 178	17 505	22	34	49
3	267	0.16604	34 695	16 489	23	36	48
4	318	0 19776	34 05	17.602	20	36	49
5	317	0.19714	33 634	16 905	22	36	48
6	186	0 11567	35 247	16 031	24	36	48
7	269	0 16729	35 093	16.104	24	36	48
8	33	0 020522	37 679	16 551	28	40	48

MENTAL CATEGORY VS. SERVICE,GED ONLY

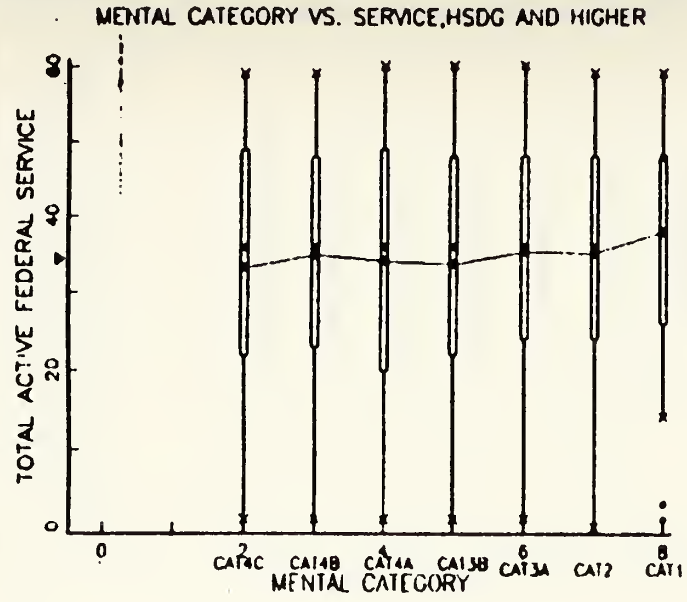


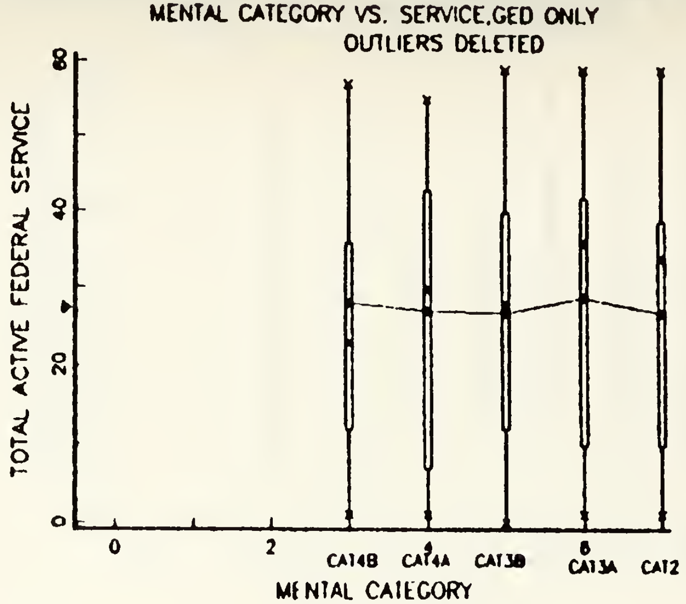
CATEGORY VECTOR :MENC1
Y :Y←A79MOD[;1]
SELECTION :MCSGED

X LABEL :ME
Y LABEL :TO
NO. OF ELEMENTS :18

CATEG.	NO PTS	o/o-PTS	Y MEAN	Y-DVN	.25	.50	.75
ALL	196	1	27 173	18.158	10	30	40
2	2	0 010204	7	4	3	3	11
3	24	0 12245	28 083	17 122	12	23	38
4	45	0 22959	27 089	18 637	7	30	43
5	57	0 29082	26 825	17 832	18	28	40
6	28	0 14286	28 857	18 15	10	38	42
7	37	0 18878	26 919	18 303	10	34	39
8	3	0.015306	28.887	20 499	1	35	50

Figure H.5 Mental Category vs. Length of Service,V





CATEGORY VECTOR :MENC1
Y :Y←A79MOD[;1]
SELECTION :MCS

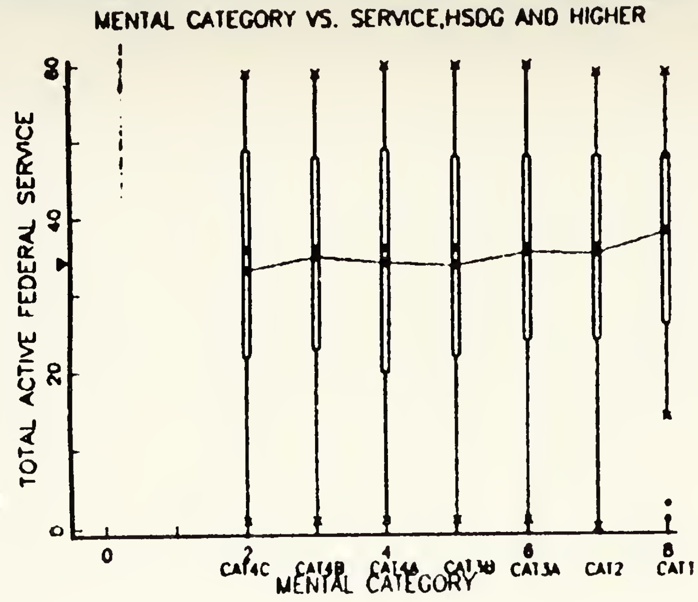
X LABEL :MF
Y LABEL :10
NO. OF ELEMENTS :16

CATEG	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1608	1	34 382	16 847	23	38	48
2	218	0 13557	33 179	17 505	22	36	49
3	287	0 16004	34.095	16 489	23	36	48
4	318	0 19776	34 05	17 602	20	36	49
5	317	0 19714	33 634	16.905	22	34	48
6	186	0 11567	35 247	16 031	24	38	48
7	269	0.16729	35 093	16 104	24	36	48
8	33	0.020522	37 879	16.551	28	48	48

CATEGORY VECTOR MENC1
Y :Y←A79MOD[;1]
SELECTION :MCSGED∧(B79MOD[;8]≥2)∧(B79MOD[;8]≤8)

CATEG	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	191	1	27 361	18.091	11	30	40
3	24	0 12565	28 083	17.122	12	23	36
4	45	0.2356	27 089	18 637	7	30	43
5	57	0.29843	26 825	17 832	12	28	40
6	28	0.1466	28 857	18 15	10	36	42
7	37	0.19372	26 919	18 303	10	34	38

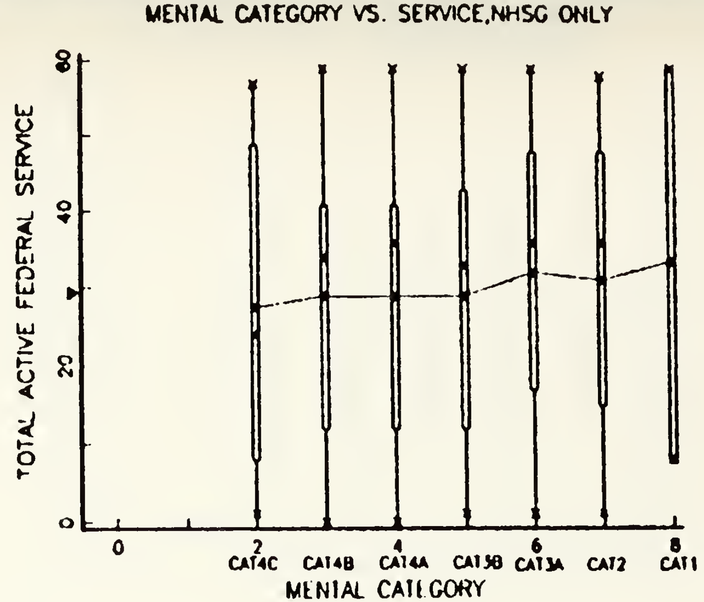
Figure H.6 Mental Category vs. Length of Service,VI



CATEGORY VECTOR :MENCT
Y :Y←A79MOD[.1]
SELECTION :MCS

X LABEL :ME
Y LABEL .10
NO. OF ELEMENTS :18

CATEG	NO PTS	0/0-PTS	Y-MEAN	Y-DVN	.25	50	75
ALL	1608	1	34 382	16 847	23	36	48
2	218	0.13557	33 179	17.505	22	36	48
3	267	0 16604	34.885	16 489	23	36	48
4	318	0 19776	34 05	17 602	20	38	49
5	317	0 19714	33.634	16 905	22	38	48
6	186	0 11567	35 247	16 031	24	38	48
7	289	0 16729	35 093	18.104	24	38	48
8	33	0.020522	37 879	16 551	26	48	48

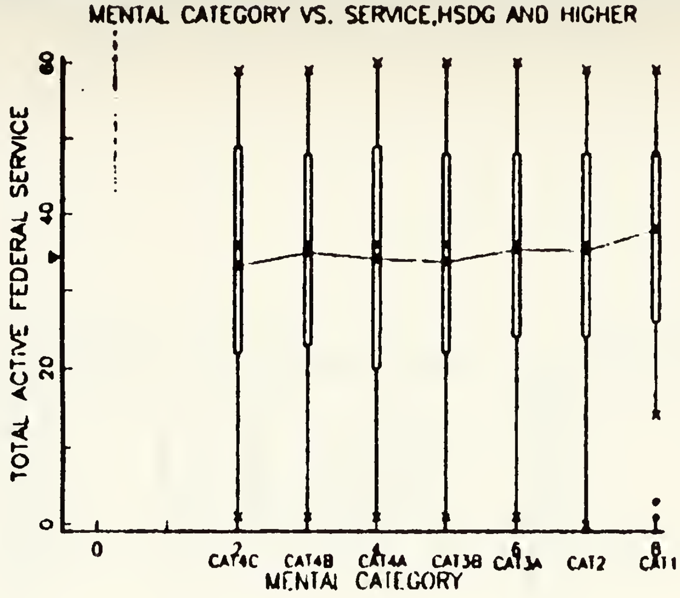


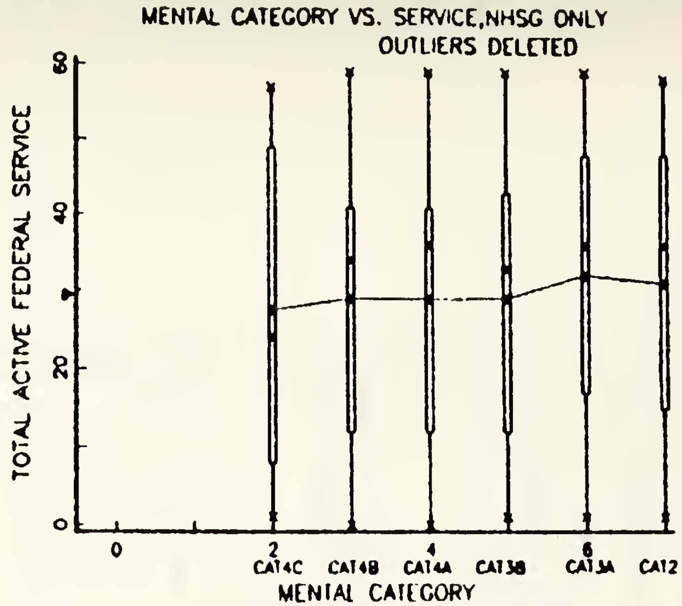
CATEGORY VECTOR :MENCT
Y :Y←A79MOD[;1]
SELECTION :MCSNHS

X LABEL :MEN
Y LABEL :TOT
NO. OF ELEMENTS :128

CATEG	NO.PTS	0/0-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1284	1	29 388	17 745	12	38	43
2	10	0 0079114	27 5	20 48	8	24	49
3	322	0 25475	28 909	17 576	12	34	41
4	433	0 34256	28 975	17 912	12	38	41
5	304	0 24051	29 003	17 701	12	33	43
6	115	0 090981	32 104	17 354	17	36	48
7	78	0 061709	31 051	17 148	15	36	48
8	2	0.0015823	33 5	25.5	8	8	59

Figure H.7 Mental Category vs. Length of Service,VII





CATEGORY VECTOR :MENCT
Y :Y←A79MOD[;1]
SELECTION :MCS

X LABEL :MI
Y LABEL :TO
NO. OF ELEMENTS :16

CATEG.	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	50	.75
ALL	1608	1	34 382	16 847	23	36	48
2	218	0.13557	33 179	17 505	22	36	48
3	267	0 16604	34 895	16 489	23	36	48
4	318	0.19776	34 05	17 602	20	36	48
5	317	0 19714	33 634	16 905	22	36	48
6	186	0 11567	35.247	16 031	24	36	48
7	269	0 16729	35 093	16 104	24	38	48
8	33	0.020522	37 879	16.551	26	48	48

CATEGORY VECTOR :MENCT
Y :Y←A79MOD[;1]
SELECTION :MCSNHS∧(B79MOD[;8]=8)

X LABEL
Y LABEL
NO. OF ELEM

CATEG.	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	50	.75
ALL	1262	1	29 382	17.73	13	36	43
2	10	0 0079239	27 5	20 40	8	24	49
3	322	0.25515	28 969	17 576	12	34	41
4	433	0 34311	28 975	17 812	12	36	41
5	304	0 24089	28.003	17 701	12	33	43
6	115	0 091125	32 104	17 354	17	36	48
7	78	0 061807	31 051	17.148	15	36	48

Figure H.8 Mental Category vs. Length of Service,VIII

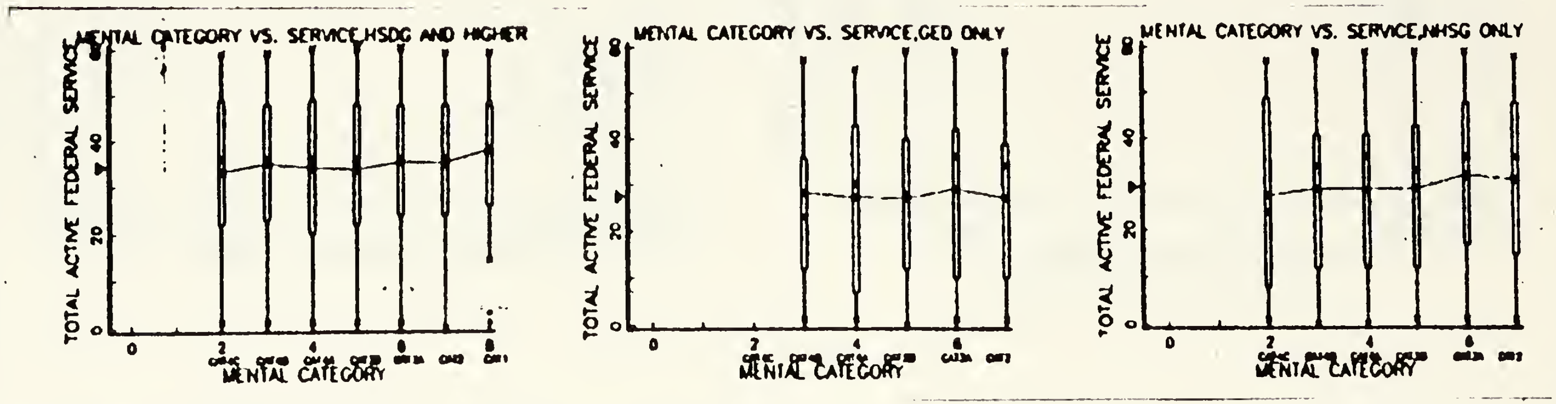
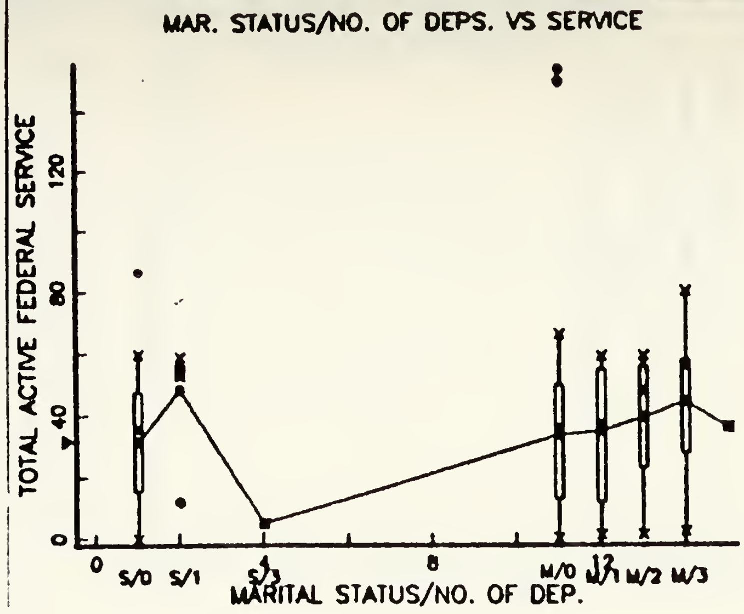
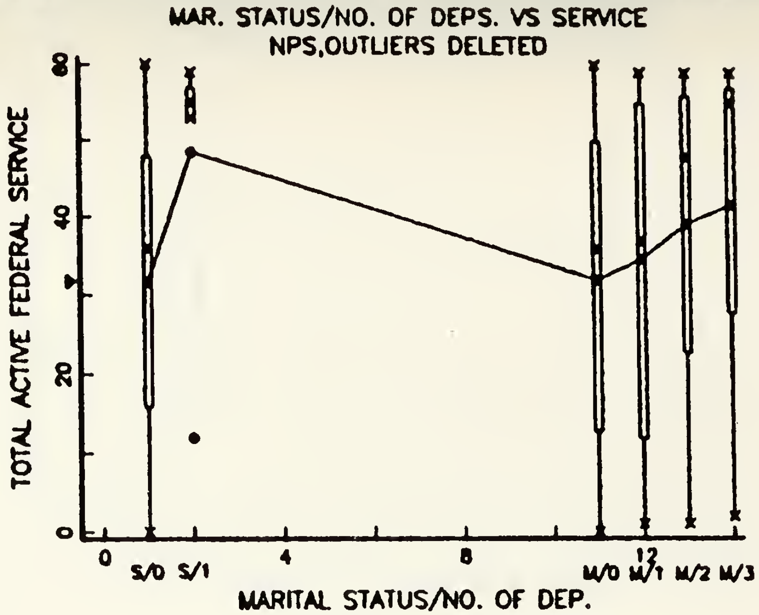


Figure H.9 Mental Category vs. Length of Service,IX





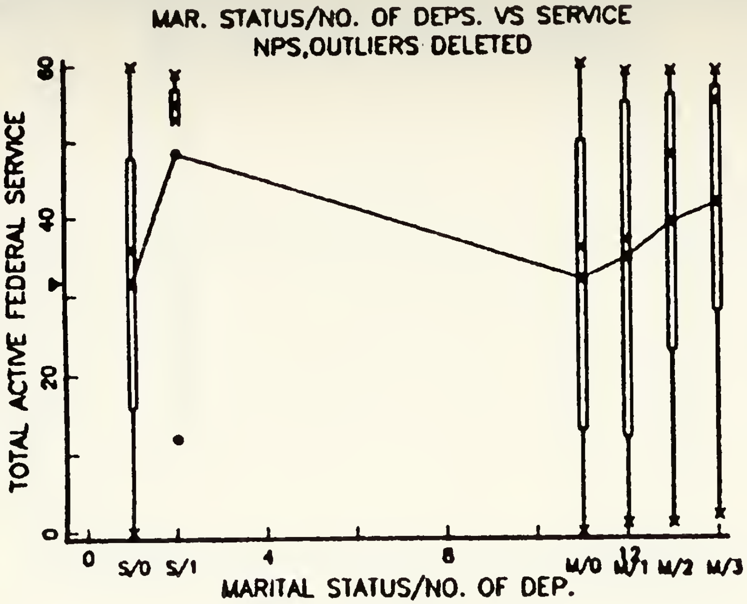
CATEGORY VECTOR :MARS   X LABEL :MARI
Y :Y←A79MOD[;1]   Y LABEL :TOTA
SELECTION :ALL   NO. OF ELEMENTS :3078

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	31.875	17.838	16	36	48
1	2843	0.92365	31.702	17.342	16	36	48
2	6	0.0019493	48.5	16.429	53	55	57
4	1	0.00032489	5	0	5	5	5
11	167	0.054256	33.641	23.131	13	38	50
12	20	0.0064977	34.7	20.737	12	37	55
13	25	0.0081222	39.2	20.781	23	48	58
14	15	0.0048733	44.267	21.874	28	56	58
15	1	0.00032489	36	0	36	36	36

CATEGORY VECTOR :MARS
Y :Y←A79MOD[;1]
SELECTION :(Y<80)^((B79MOD[;5]≠13)^(B79MOD[;5]≠24))

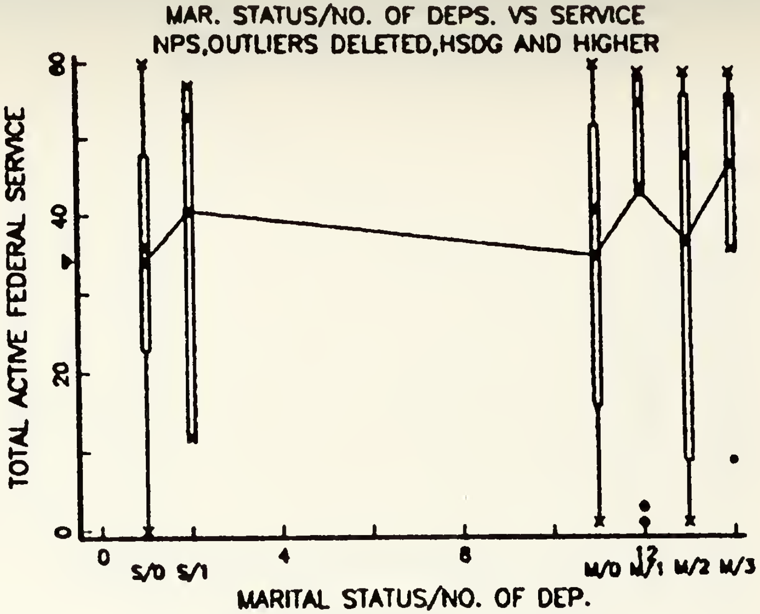
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3071	1	31.858	17.527	16	36	48
1	2842	0.92543	31.683	17.314	16	36	48
2	6	0.0019538	48.5	16.429	53	55	57
11	164	0.053403	32	18.101	13	36	50
12	20	0.0065125	34.7	20.737	12	37	55
13	25	0.0081407	39.2	20.781	23	48	56
14	14	0.0045588	41.714	20.485	28	55	57

Figure H.10 Marital Status vs. Service,I



CATEGORY VECTOR :MARS
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧((B79MOD[;5]≠13)∧(B79MOD[;5]≠24))

CATEG.	NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3071	1	31.858	17.527	16	36	48
1	2842	0.92543	31.683	17.314	16	36	48
2	6	0.0019538	48.5	16.429	53	55	57
11	164	0.053403	32	19.101	13	36	50
12	20	0.0065125	34.7	20.737	12	37	55
13	25	0.0081407	39.2	20.781	23	48	58
14	14	0.0045588	41.714	20.485	28	55	57

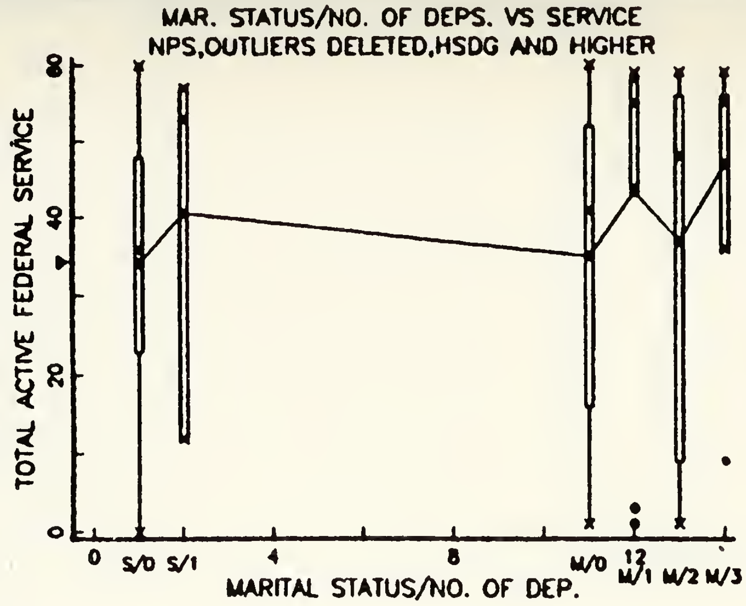


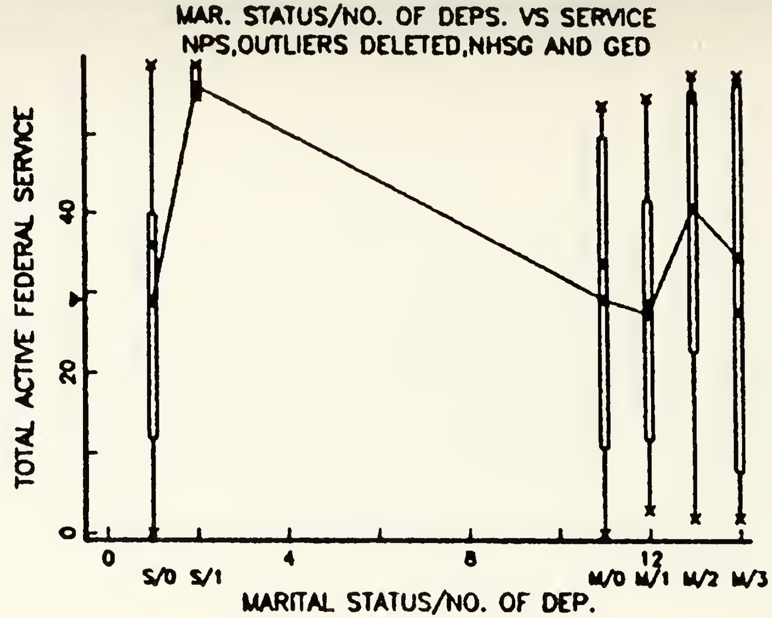
CATEGORY VECTOR :MARS
Y :Y←A79MOD[;1]
SELECTION :SNHS

X LABEL :MAR
Y LABEL :TOT
NO. OF ELEMENTS :181

CATEG.	NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.849	22	36	48
1	1505	0.93362	34.183	16.651	23	36	48
2	3	0.001861	40.667	20.336	12	53	57
11	76	0.047146	35.013	16.544	18	41	52
12	9	0.0055831	43.333	22.509	44	55	58
13	11	0.0068238	36.909	21.753	9	48	58
14	8	0.0049628	46.875	16.035	38	55	58

Figure H.11 Marital Status vs. Service,II





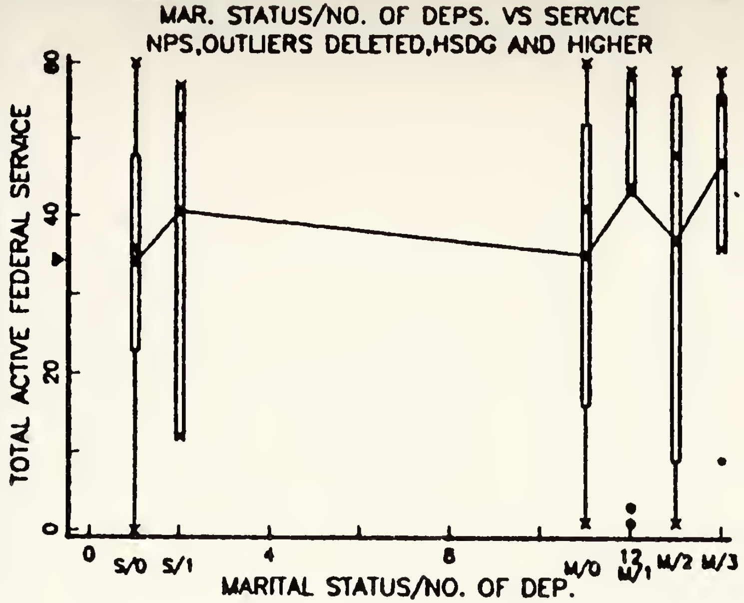
CATEGORY VECTOR :MARS X LABEL :MAR
Y :Y-A79MOD[;1] Y LABEL :TOT
SELECTION :SNHS NO. OF ELEMENTS :161

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.347	16.869	22	36	48
1	1505	0.93362	34.183	16.651	23	36	48
2	3	0.001861	40.667	20.336	12	53	57
11	76	0.047146	35.013	16.544	16	41	52
12	9	0.0055831	43.333	22.509	44	55	58
13	11	0.0068238	36.909	21.753	9	48	58
14	8	0.0049628	46.875	16.035	36	55	58

CATEGORY VECTOR :MARS X LABEL :MAR
Y :Y-A79MOD[;1] Y LABEL :TOT
SELECTION :SNNHS NO. OF ELEMENTS :145

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1459	1	29.008	17.82	12	36	43
1	1337	0.91638	28.868	17.613	12	36	40
2	3	0.0020562	56.333	1.8856	55	55	58
11	88	0.060315	29.398	19.183	11	34	50
12	11	0.0075394	27.636	16.013	12	29	42
13	14	0.0095956	41	19.799	23	55	56
14	6	0.0041124	34.833	23.526	8	28	57

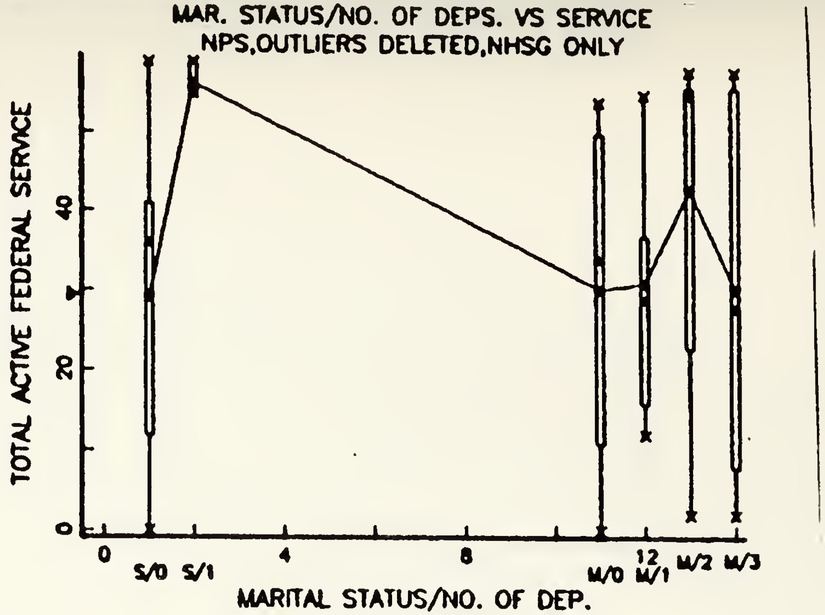
Figure B.12 Marital Status vs. Service,III



CATEGORY VECTOR :MARS
Y :Y←A79MOD[;1]
SELECTION :SNHS

X LABEL :MAR
Y LABEL :TOT
NO. OF ELEMENTS :161

CATEG.	NO.PTS	●/●-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	38	48
1	1505	0.93362	34.183	16.651	23	38	48
2	3	0.001861	40.667	20.336	12	53	57
11	76	0.047146	35.013	18.544	18	41	52
12	9	0.0055831	43.333	22.509	44	55	58
13	11	0.0068238	36.909	21.753	9	48	56
14	8	0.0049628	46.875	16.035	36	55	56

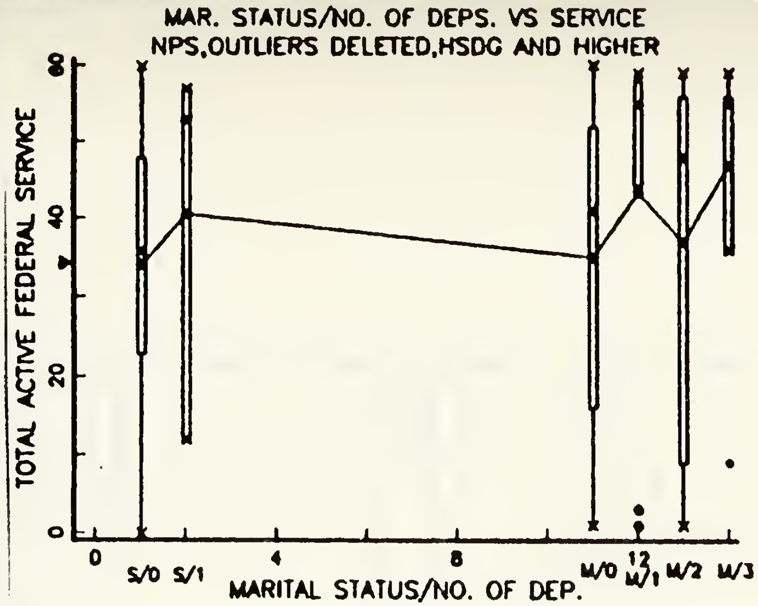


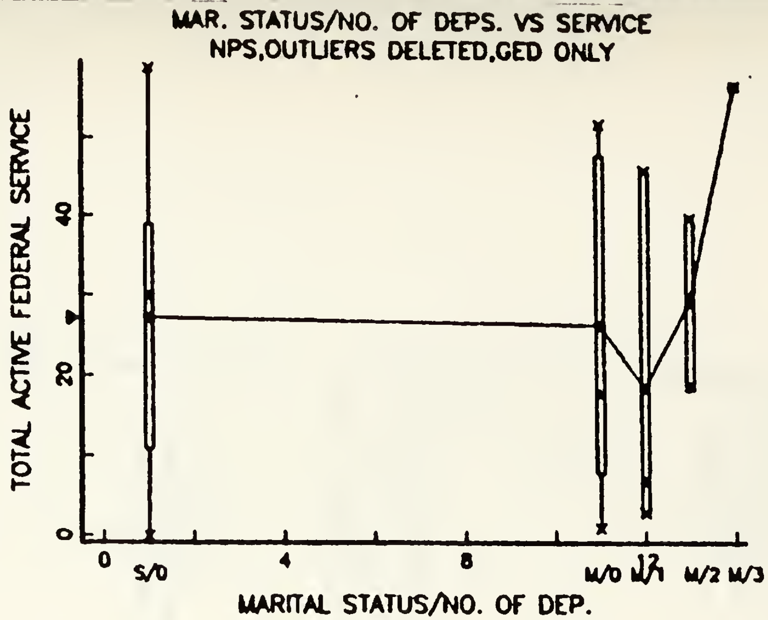
CATEGORY VECTOR :MARS
Y :Y←A79MOD[;1]
SELECTION :SNHSG

X LABEL :MAR
Y LABEL :TOT
NO. OF ELEMENTS :128

CATEG.	NO.PTS	●/●-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1263	1	29.386	17.748	12	36	44
1	1167	0.92399	29.111	17.58	12	36	41
2	3	0.0023753	56.333	1.8856	55	55	59
11	68	0.05384	30.25	16.938	11	34	50
12	8	0.0063341	31	13.038	16	29	37
13	12	0.0095012	42.917	20.328	23	55	56
14	5	0.0039588	30.4	23.372	8	28	58

Figure H.13 Marital Status vs. Service,IV





CATEGORY VECTOR :MARS
Y :Y-A79MOD[;1]
SELECTION :SNHS

X LABEL :MAR
Y LABEL :TOT
NO. OF ELEMENTS :161

CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	38	48
1	1505	0.93362	34.183	16.651	23	38	48
2	3	0.001861	40.667	20.336	12	53	57
11	76	0.047146	35.013	18.544	18	41	52
12	9	0.0055831	43.333	22.509	44	55	58
13	11	0.0068238	38.909	21.753	9	48	58
14	8	0.0049628	46.875	16.035	38	55	56

CATEGORY VECTOR :MARS
Y :Y-A79MOD[;1]
SELECTION :SNGED

X LABEL :MA
Y LABEL :TO
NO. OF ELEMENTS :19

CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	196	1	27.173	18.150	10	30	40
1	170	0.86735	27.2	17.879	11	30	38
11	20	0.10204	26.5	19.782	8	18	48
12	3	0.015306	18.667	19.346	3	7	46
13	2	0.010204	28.5	10.5	18	18	40
14	1	0.005102	57	0	57	57	57

Figure H.14 Marital Status vs. Service,V

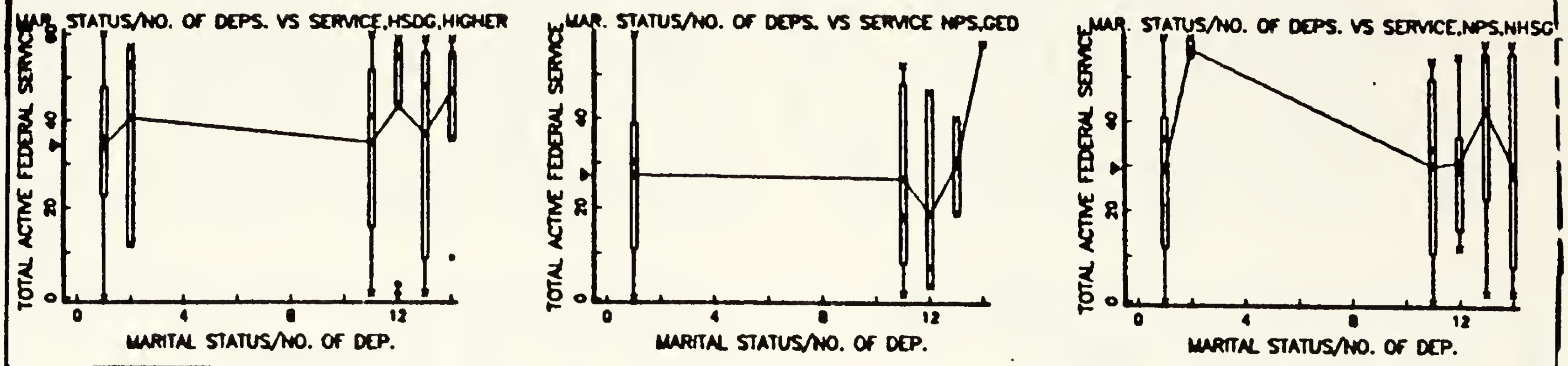
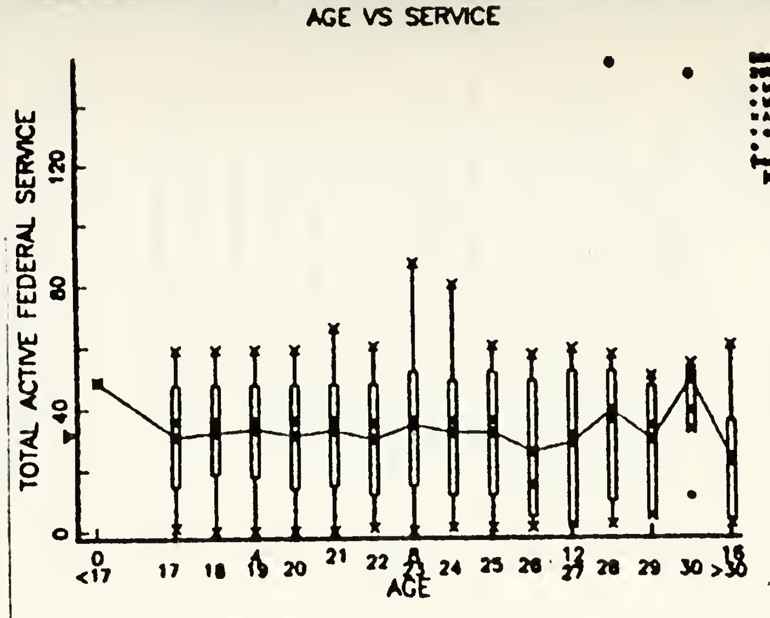
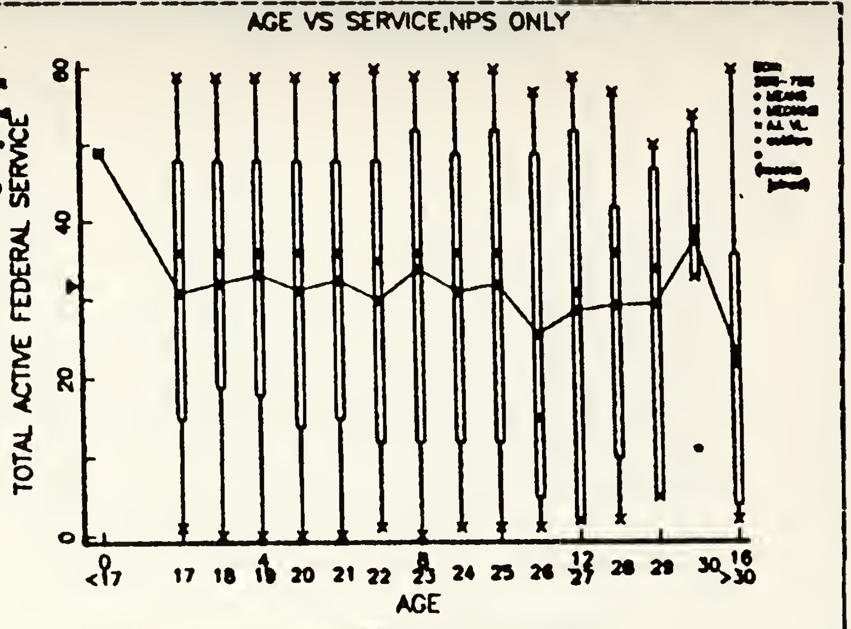


Figure H.15 Marital Status vs. Service,VI

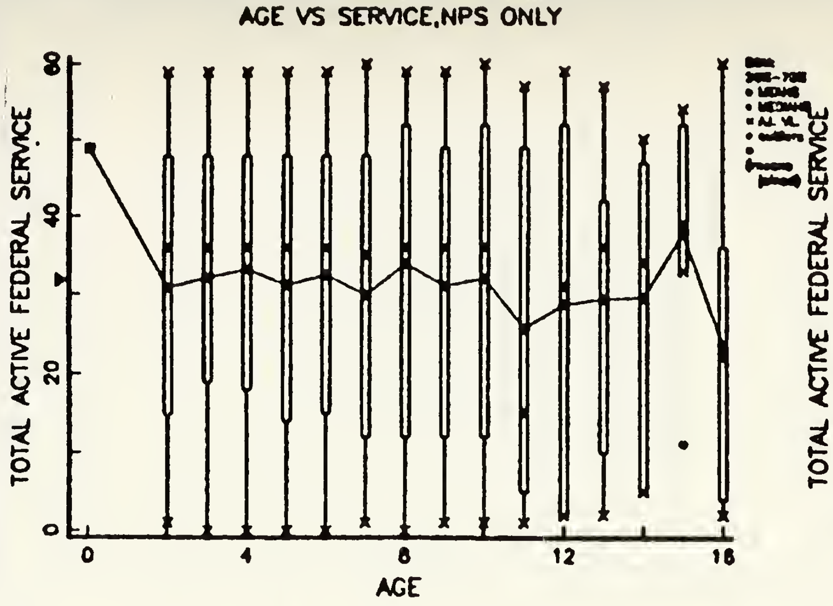


CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	31.875	17.638	18	36	48
0	1	0.00032489	49	0	49	49	49
2	330	0.10721	30.855	16.517	15	36	48
3	1132	0.36777	32.099	16.672	19	36	48
4	650	0.21118	33.152	17.603	18	36	48
5	338	0.10981	31.142	16.043	14	36	48
6	183	0.059454	32.617	17.934	15	36	48
7	134	0.043535	29.896	16.347	12	35	48
8	85	0.027615	34.518	19.96	15	36	52
9	55	0.017869	31.927	20.17	12	36	49
10	47	0.01527	31.936	19.752	12	36	52
11	38	0.012346	25.605	21.168	5	15	49
12	27	0.0087719	28.63	21.408	2	31	52
13	13	0.0042235	38.923	38.418	10	36	52
14	8	0.0025991	29.5	18.808	5	34	47
15	10	0.0032489	48.9	38.863	33	39	52
16	27	0.0087719	23.444	18.612	4	22	36

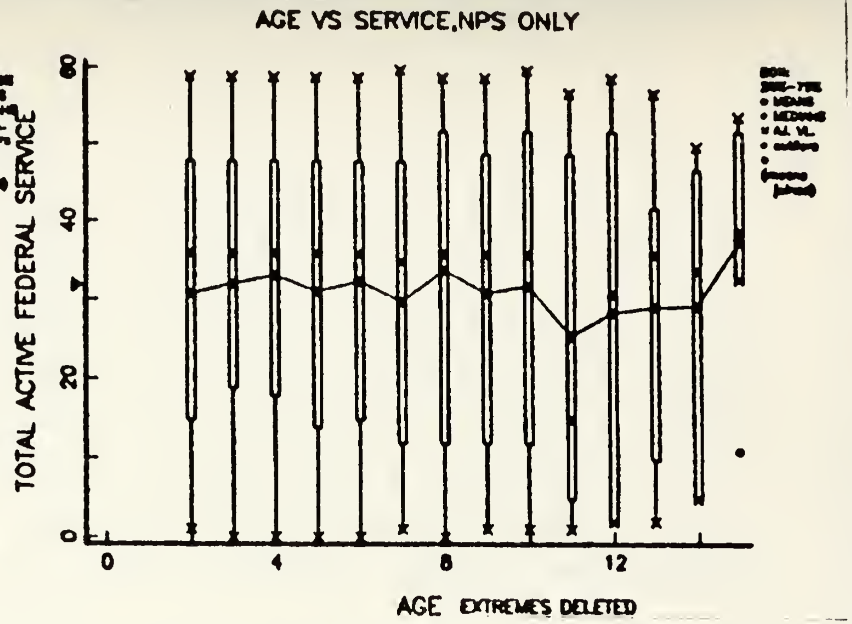


CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	36	48
0	1	0.00032541	49	0	49	49	49
2	330	0.10739	30.855	16.517	15	36	48
3	1132	0.36837	32.099	16.672	19	36	48
4	650	0.21152	33.152	17.603	18	36	48
5	338	0.10999	31.142	16.043	14	36	48
6	182	0.059226	32.434	17.811	15	36	48
7	134	0.043606	29.896	16.347	12	35	48
8	84	0.027335	33.893	19.234	12	36	52
9	54	0.017572	31.037	19.258	12	36	49
10	47	0.015295	31.936	19.752	12	36	52
11	38	0.012366	25.605	21.168	5	15	49
12	27	0.0087862	28.63	21.408	2	31	52
13	12	0.003905	29.333	20.085	10	36	42
14	8	0.0026033	29.5	18.808	5	34	47
15	9	0.0029287	37.667	15.748	33	39	52
16	27	0.0087862	23.444	18.612	4	22	36

Figure H.16   Age vs, Service,I

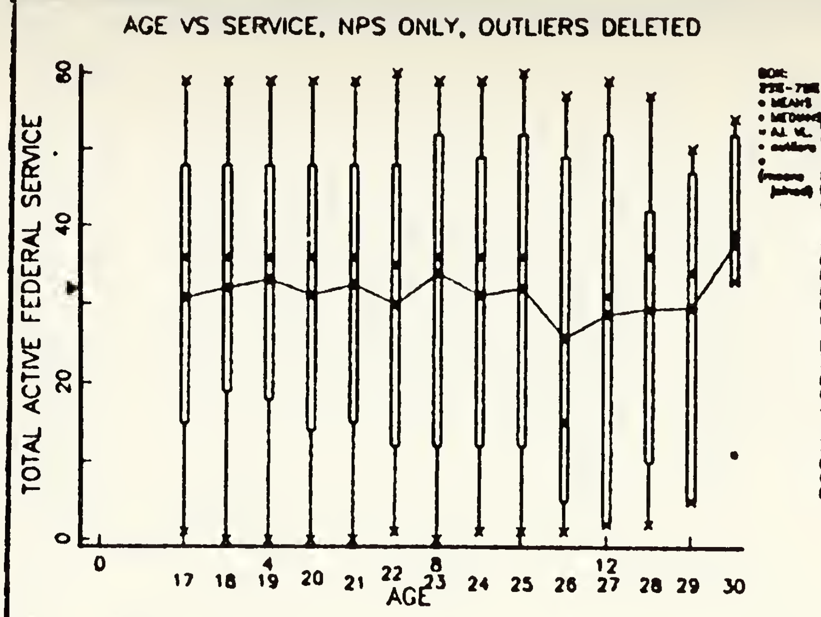


CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	36	48
0	1	0.00032541	49	0	49	49	49
2	330	0.10739	30.855	18.517	15	36	48
3	1132	0.36837	32.099	18.672	19	36	48
4	650	0.21152	33.152	17.803	18	36	48
5	338	0.10999	31.142	18.043	14	36	48
6	182	0.059228	32.434	17.811	15	36	48
7	134	0.043606	29.896	18.347	12	35	48
8	84	0.027335	33.893	19.234	12	36	52
9	54	0.017572	31.037	18.256	12	36	49
10	47	0.015295	31.936	19.752	12	36	52
11	38	0.012366	25.605	21.168	5	15	49
12	27	0.0087862	28.63	21.408	2	31	52
13	12	0.003905	29.333	20.065	10	36	42
14	8	0.0026033	29.5	18.808	5	34	47
15	8	0.0029287	37.667	15.748	33	39	52
16	27	0.0087862	23.444	18.812	4	22	36

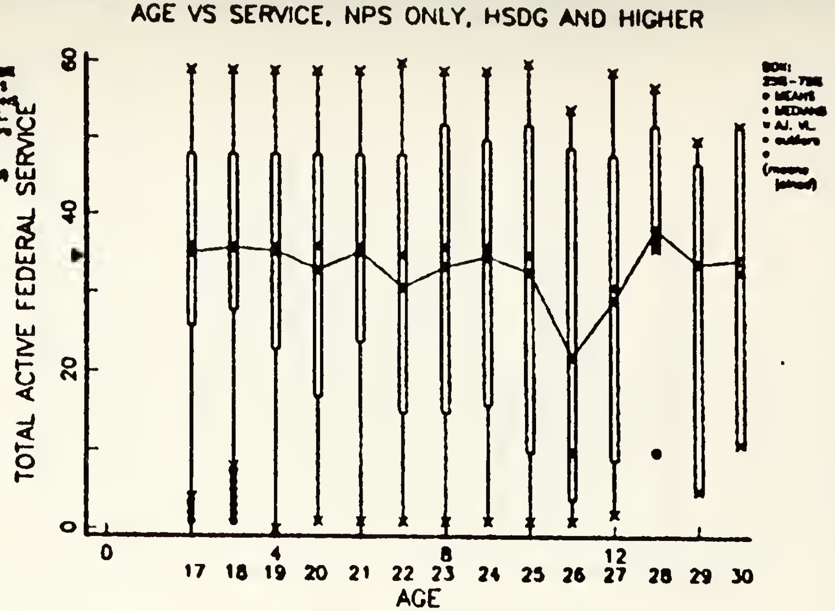


CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3045	1	31.921	17.5	18	36	48
2	330	0.10837	30.855	18.517	15	36	48
3	1132	0.37176	32.099	18.672	18	36	48
4	650	0.21346	33.152	17.803	18	36	48
5	338	0.111	31.142	18.043	14	36	48
6	182	0.05977	32.434	17.811	15	36	48
7	134	0.044007	29.896	18.347	12	35	48
8	84	0.027586	33.893	19.234	12	36	52
9	54	0.017734	31.037	19.256	12	36	48
10	47	0.015435	31.936	19.752	12	36	52
11	38	0.012479	25.605	21.168	5	15	49
12	27	0.008867	28.63	21.406	2	31	52
13	12	0.0039409	29.333	20.065	10	36	42
14	8	0.0026273	29.5	18.808	5	34	47
15	9	0.0029557	37.667	15.748	33	39	52

Figure H.17 Age vs. Service,II

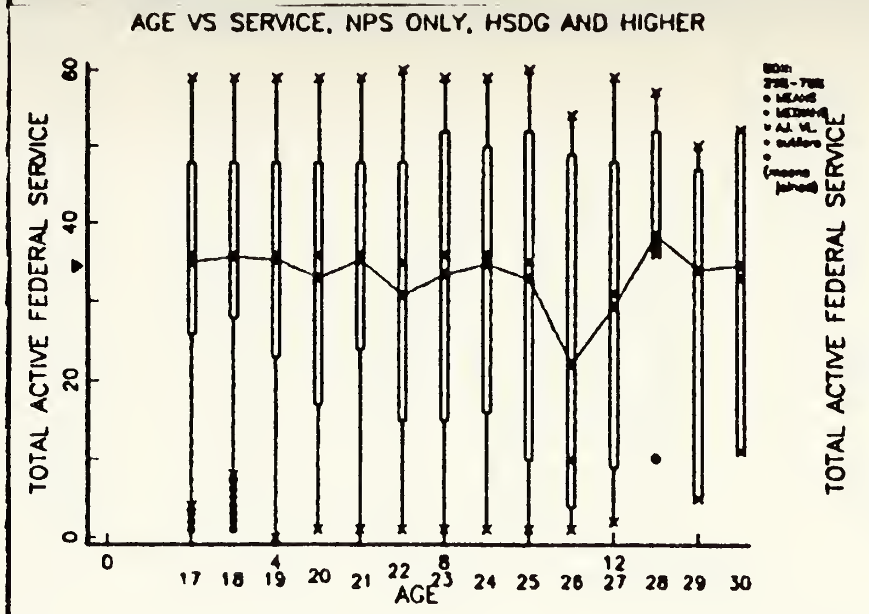


CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3045	1	31.921	17.5	16	36	48
2	330	0.10837	30.855	16.517	15	36	48
3	1132	0.37176	32.099	16.672	19	36	48
4	650	0.21346	33.152	17.603	18	36	48
5	338	0.111	31.142	18.043	14	36	48
6	182	0.05977	32.434	17.811	15	36	48
7	134	0.044007	29.896	18.347	12	35	48
8	84	0.027586	33.893	19.234	12	36	52
9	54	0.017734	31.037	19.258	12	36	49
10	47	0.015435	31.936	19.752	12	36	52
11	38	0.012479	25.605	21.168	5	15	49
12	27	0.008867	28.63	21.408	2	31	52
13	12	0.0039409	29.333	20.085	10	36	42
14	8	0.0026273	29.5	18.808	5	34	47
15	9	0.0029557	37.667	15.748	33	39	52

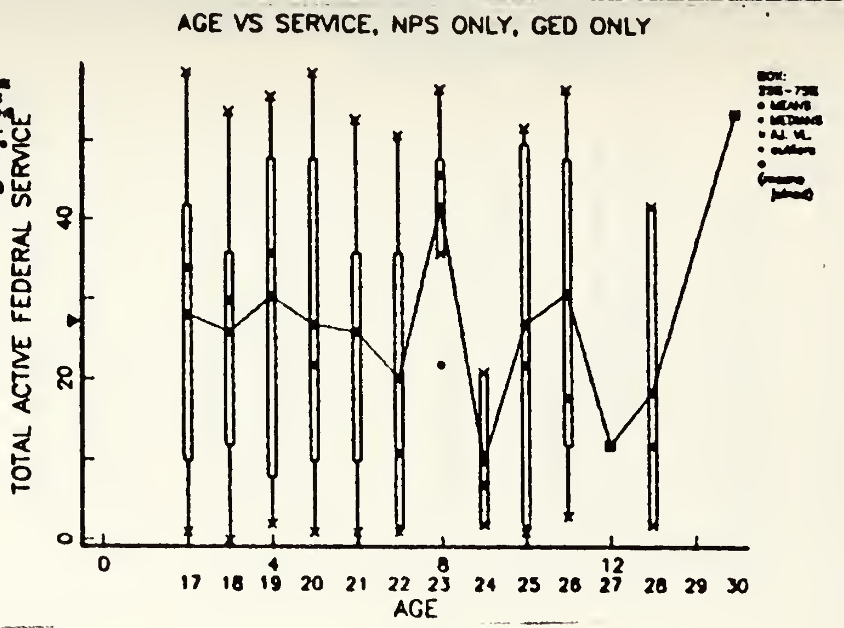


CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1593	1	34.513	16.785	23	36	48
2	149	0.093534	35.054	15.186	26	36	48
3	527	0.33082	35.898	15.327	28	36	48
4	356	0.22348	35.343	16.901	23	36	48
5	200	0.12555	33 015	17.74	17	36	48
6	101	0.063402	35.307	16.539	24	36	48
7	83	0.052103	30.783	16.935	15	35	48
8	53	0.033271	33.472	16.882	15	36	52
9	34	0.021343	34.676	19.525	16	36	50
10	32	0.020088	32.844	19.867	10	35	52
11	22	0.01381	22	20.445	4	10	49
12	21	0.013183	29.286	21.189	9	31	48
13	5	0.0031387	38.4	16.402	38	37	52
14	4	0.002511	34	17.79	5	34	47
15	6	0.0037665	34.5	17.802	11	33	52

Figure H.18 Age vs. Service,III

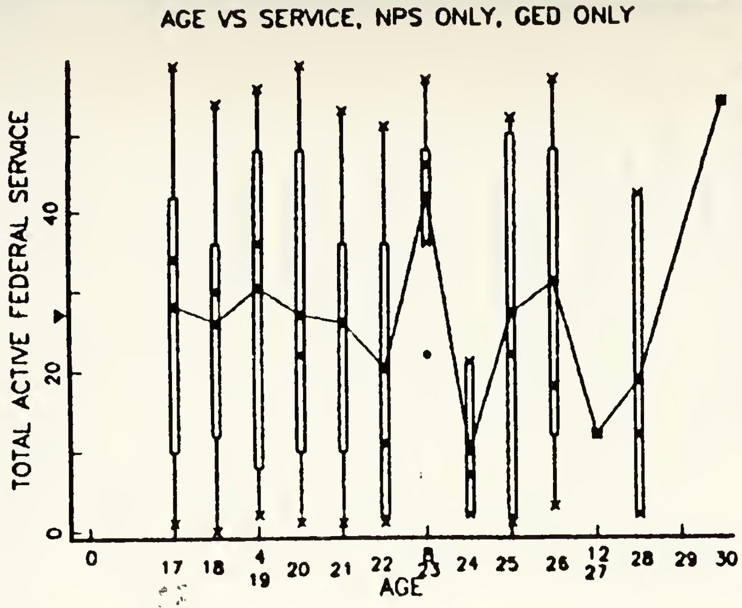


CATEG.	NO.PTS	I•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1593	1	34.513	16.785	23	38	48
2	149	0.093534	35.054	15.186	28	36	48
3	527	0.33082	35.696	15.327	28	36	48
4	356	0.22348	35.343	16.901	23	36	48
5	200	0.12555	33.015	17.74	17	36	48
6	101	0.063402	35.307	16.539	24	36	48
7	83	0.052103	30.783	16.935	15	35	48
8	53	0.033271	33.472	18.882	15	36	52
9	34	0.021343	34.676	19.525	16	36	50
10	32	0.020088	32.844	19.867	10	35	52
11	22	0.01381	22	20.445	4	10	49
12	21	0.013183	29.286	21.169	9	31	48
13	5	0.0031387	38.4	16.402	36	37	52
14	4	0.002511	34	17.79	5	34	47
15	6	0.0037665	34.5	17.802	11	33	52

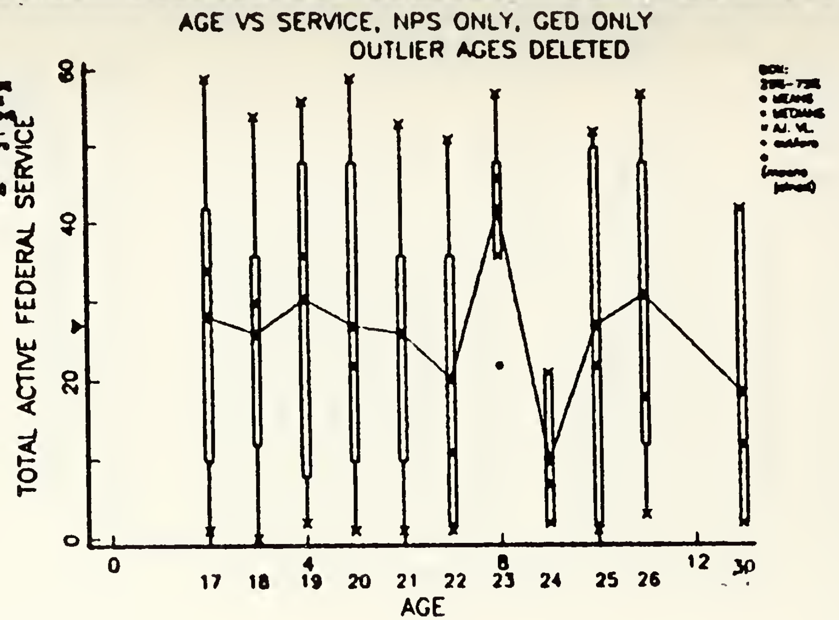


CATEG.	NO.PTS	I•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	195	1	27.18	18.204	10	30	40
2	23	0.11795	28.13	18.613	10	34	42
3	66	0.33846	28.015	16.049	12	30	36
4	35	0.17949	30.429	18.721	8	34	48
5	23	0.11795	27	19.222	10	22	48
6	13	0.066667	26.077	18.257	10	26	36
7	10	0.051282	20.3	18.639	2	11	36
8	5	0.025641	41.8	11.94	36	46	48
9	3	0.015385	10	8.0416	2	7	21
10	6	0.030769	27.167	20.659	2	22	50
11	6	0.030769	31	20.888	12	18	48
12	1	0.0051282	12	0	12	12	12
13	3	0.015385	18.667	16.997	2	12	42
15	1	0.0051282	54	0	54	54	54

Figure H.19 Age vs. Service,IV

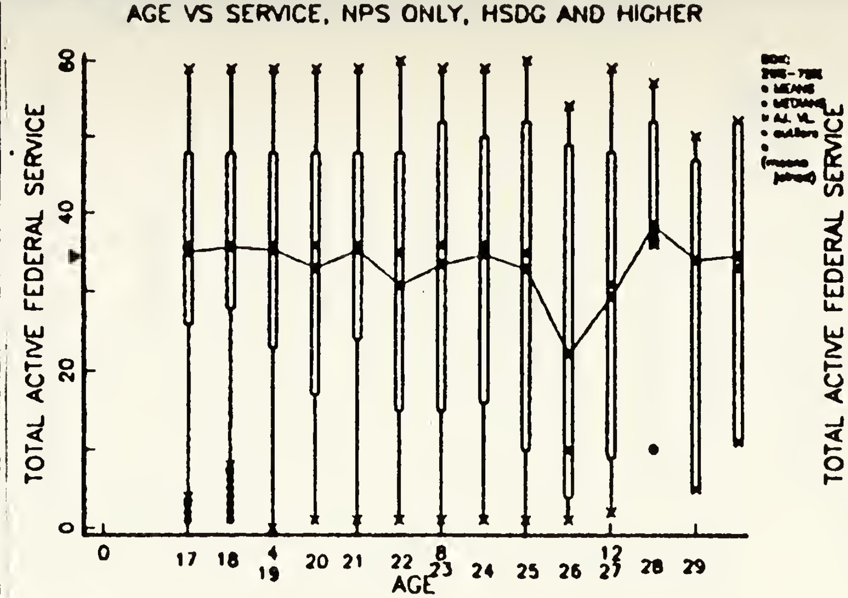


CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	195	1	27.19	18.204	10	30	40
2	23	0.11795	28.13	18.613	10	34	42
3	66	0.33846	26.015	16.049	12	30	36
4	35	0.17949	30.429	18.721	8	36	48
5	23	0.11795	27	19.222	10	22	48
6	13	0.066667	26.077	18.257	10	26	36
7	10	0.051282	20.3	18.639	2	11	36
8	5	0.025641	41.8	11.94	36	46	48
9	3	0.015385	10	8.0416	2	7	21
10	6	0.030769	27.167	20.659	2	22	50
11	6	0.030769	31	20.888	12	18	48
12	1	0.0051282	12	0	12	12	12
13	3	0.015385	18.667	16.997	2	12	42
15	1	0.0051282	54	0	54	54	54

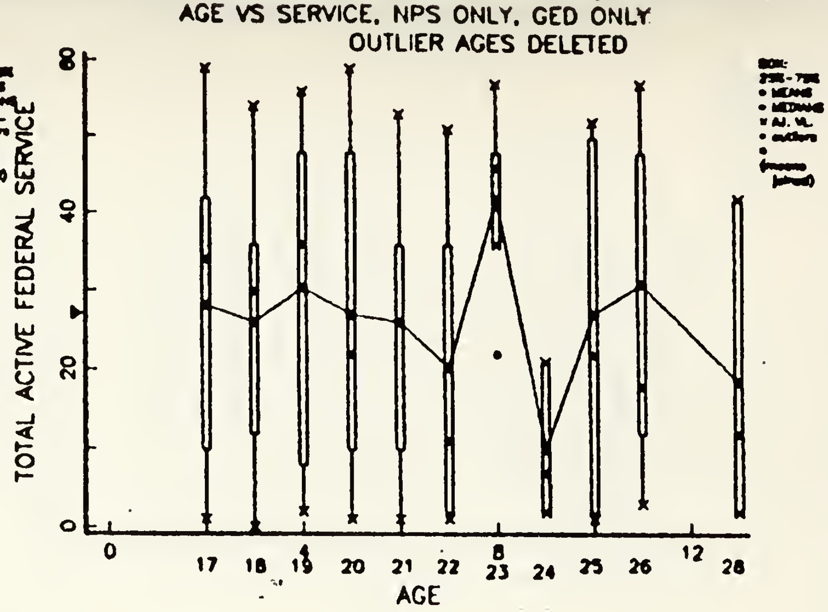


CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	193	1	27.13	18.163	10	30	40
2	23	0.11917	28.13	18.613	10	34	42
3	66	0.34197	26.015	16.049	12	30	36
4	35	0.18135	30.429	18.721	8	36	48
5	23	0.11917	27	19.222	10	22	48
6	13	0.067358	26.077	18.257	10	26	36
7	10	0.051813	20.3	18.639	2	11	36
8	5	0.025907	41.8	11.94	36	46	48
9	3	0.015544	10	8.0418	2	7	21
10	6	0.031088	27.167	20.659	2	22	50
11	6	0.031088	31	20.888	12	18	48
13	3	0.015544	18.667	16.997	2	12	42

Figure H.20 Age vs. Service,V

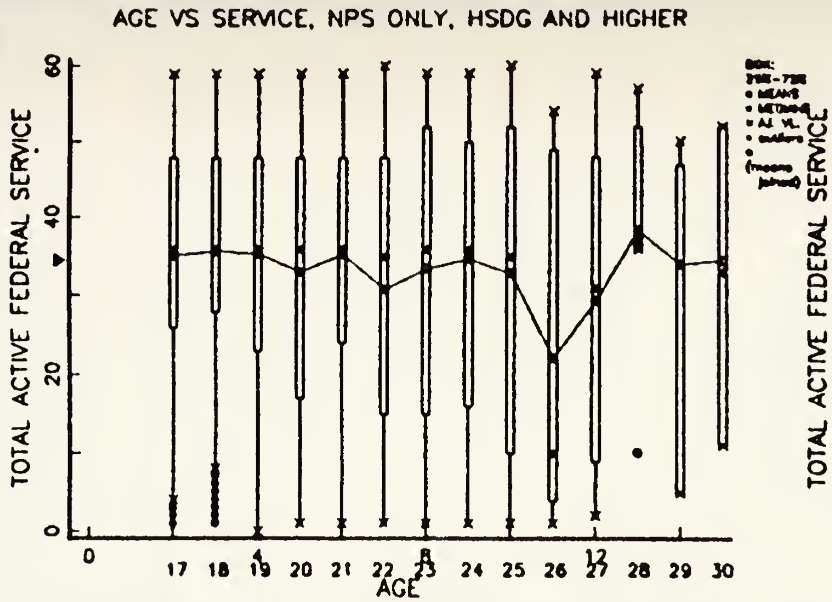


CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1593	1	34.513	16.785	23	36	48
2	149	0.093534	35.054	15 186	26	36	48
3	527	0.33082	35.696	15.327	28	36	48
4	356	0.22348	35.343	16.901	23	36	48
5	200	0.12555	33.015	17.74	17	36	48
6	101	0.063402	35.307	16 539	24	36	48
7	83	0.052103	30.783	16 935	15	35	48
8	53	0.033271	33.472	18.882	15	36	52
9	34	0.021343	34.676	19.525	16	36	50
10	32	0.020088	32.844	19.867	10	35	52
11	22	0 01381	22	20.445	4	10	49
12	21	0.013183	29.286	21.189	9	31	48
13	5	0.0031387	38.4	16.402	38	37	52
14	4	0.002511	34	17.79	5	34	47
15	6	0.0037665	34.5	17.802	11	33	52

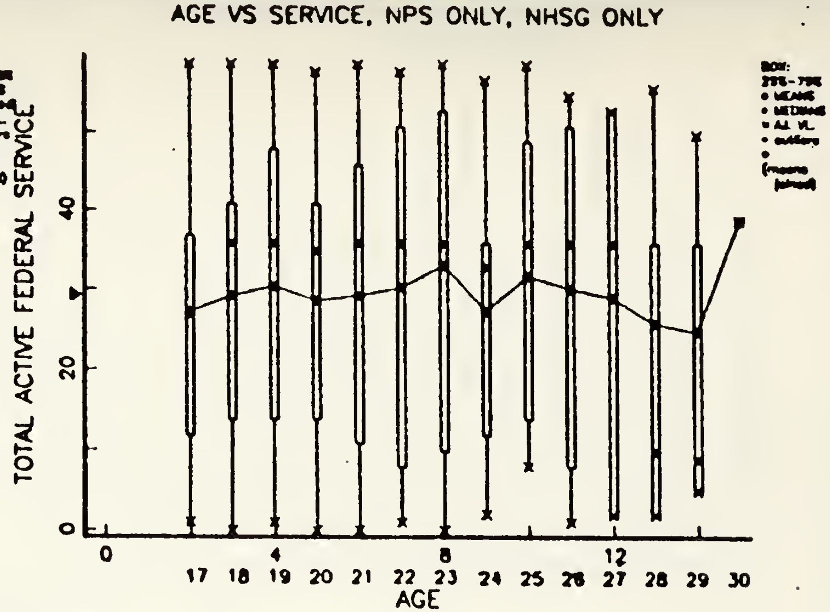


CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	193	1	27.13	18.163	10	30	40
2	23	0.11917	28.13	18.613	10	34	42
3	66	0.34197	26.015	16.049	12	30	36
4	35	0.18135	30.429	18.721	8	36	48
5	23	0.11917	27	19.222	10	22	48
6	13	0.067358	26.077	18.257	10	26	36
7	10	0.051813	20.3	18.639	2	11	36
8	5	0.025907	41.8	11.94	36	48	48
9	3	0.015544	10	8.0416	2	7	21
10	6	0.031088	27.167	20.659	2	22	50
11	6	0.031088	31	20.688	12	18	48
13	3	0.015544	18.667	16.997	2	12	42

Figure H.21 Age vs. Service,VI

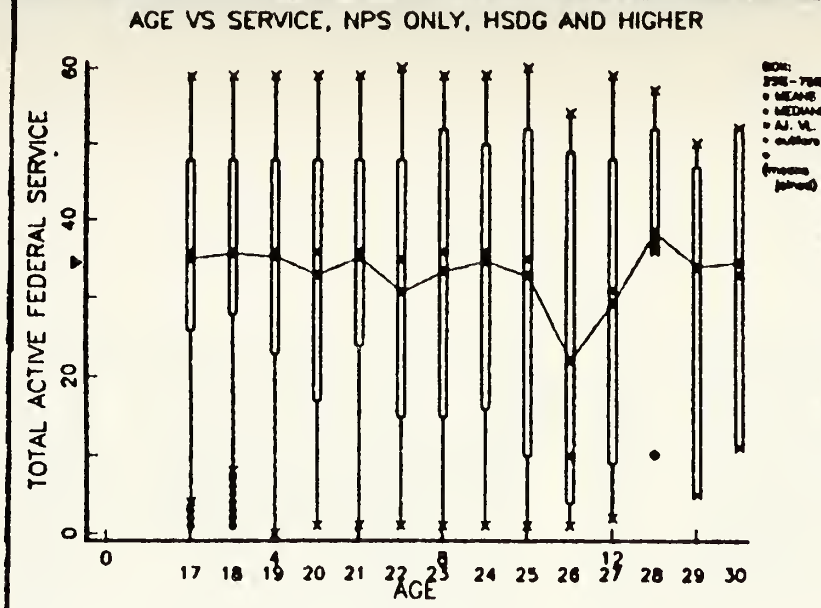


CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1593	1	34.513	16.785	23	36	48
2	149	0.093534	35.054	15.188	26	36	48
3	527	0.33082	35.696	15.327	28	36	48
4	356	0.22348	35.343	16.901	23	34	48
5	200	0.12555	33.015	17.74	17	36	48
6	101	0.063402	35.307	16.539	24	36	48
7	83	0.052103	30.783	16.935	15	35	48
8	53	0.033271	33.472	18.882	15	36	52
9	34	0.021343	34.676	19.525	16	36	50
10	32	0.020088	32.844	19.867	10	35	52
11	22	0 01381	22	20.445	4	10	49
12	21	0.013183	29.286	21.189	9	31	48
13	5	0.0031387	38 4	16.402	36	37	52
14	4	0.002511	34	17.79	5	34	47
15	6	0.0037665	34.5	17.802	11	33	52

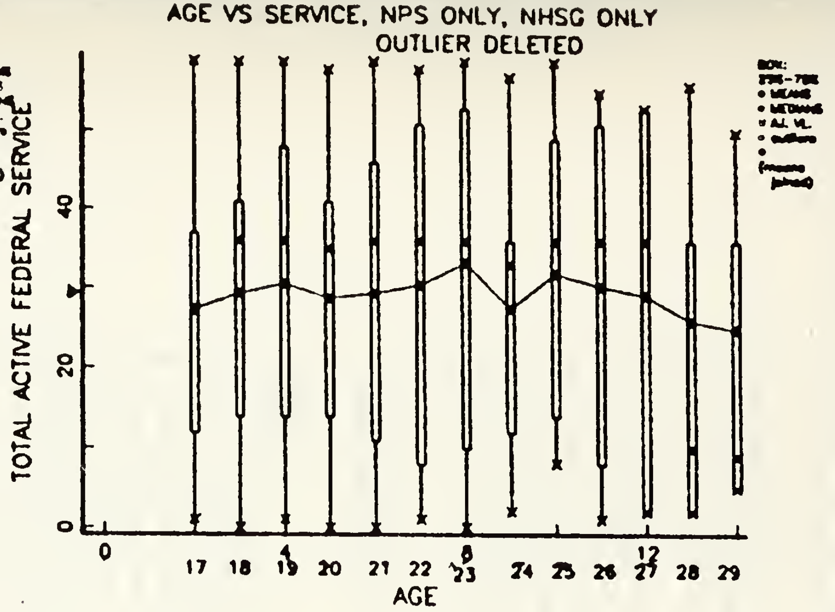


CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1257	1	29.37	17.746	12	36	43
2	158	0.1257	27.291	16.476	12	27	37
3	539	0.4288	29.327	17.254	14	36	41
4	259	0.20605	30 51	17.976	14	36	48
5	115	0.091488	28.713	17.892	14	35	41
6	68	0.054097	29.382	18.701	11	36	46
7	41	0.032617	30 439	20.261	8	36	51
8	26	0.020684	33.231	20.694	10	36	53
9	17	0.013524	27.471	16.929	12	33	36
10	9	0.0071599	31.889	18.205	14	36	49
11	10	0.0079554	30.3	21.406	8	36	51
12	5	0.0039777	29.2	23.06	2	36	53
13	4	0.0031822	26	21.401	2	10	36
14	4	0 0031822	25	18.722	5	9	36
15	2	0.0015911	39	0	39	39	39

Figure H.22 Age vs. Service,VII



CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1593	1	34.513	18.785	23	36	48
2	149	0.093534	35.054	15.186	26	36	48
3	527	0.33082	35.698	15.327	28	36	48
4	356	0.22348	35.343	16.901	23	36	48
5	200	0.12555	33.015	17.74	17	36	48
6	101	0.063402	35.307	18.539	24	36	48
7	83	0.052103	30.783	18.935	15	35	48
8	53	0.033271	33.472	18.882	15	36	52
9	34	0.021343	34.676	19.525	16	36	50
10	32	0.020088	32.844	19.867	10	35	52
11	22	0.01381	22	20.445	4	10	49
12	21	0.013183	29.286	21.189	9	31	48
13	5	0.0031387	38.4	16.402	36	37	52
14	4	0.002511	34	17.79	5	34	47
15	6	0.0037665	34.5	17.802	11	33	52



CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1255	1	29.355	17.756	12	36	44
2	158	0.1259	27.291	16.476	12	27	37
3	539	0.42948	29.327	17.254	14	36	41
4	259	0.20637	30.51	17.978	14	36	48
5	115	0.091633	28.713	17.892	14	35	41
6	68	0.054183	29.382	18.701	11	36	46
7	41	0.032669	30.439	20.261	8	36	51
8	26	0.020717	33.231	20.694	10	36	53
9	17	0.013546	27.471	16.929	12	33	36
10	9	0.0071713	31.889	18.205	14	36	49
11	10	0.0079681	30.3	21.406	8	36	51
12	5	0.0039841	29.2	23.06	2	36	53
13	4	0.0031873	26	21.401	2	10	36
14	4	0.0031873	25	18.722	5	9	36

Figure H.23 Age vs. Service,VIII

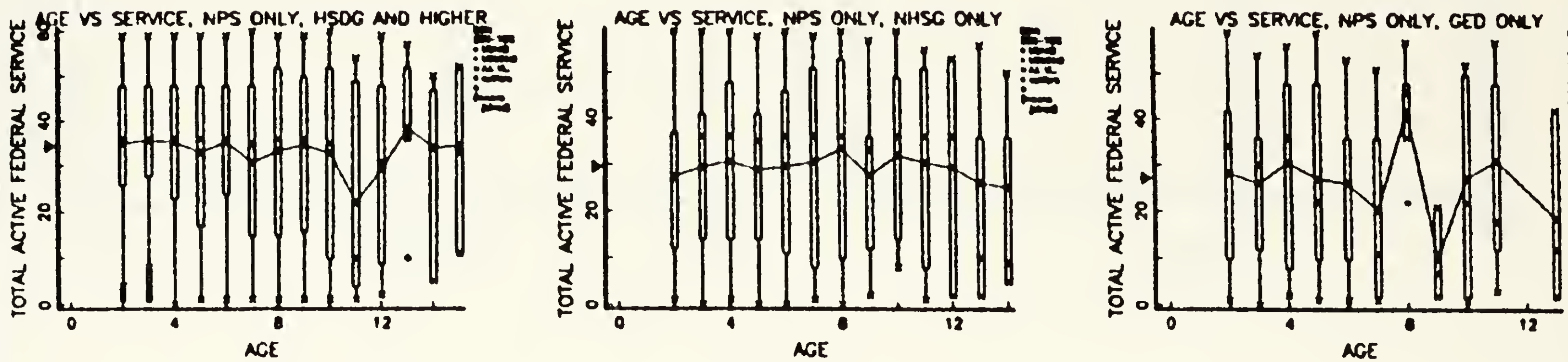
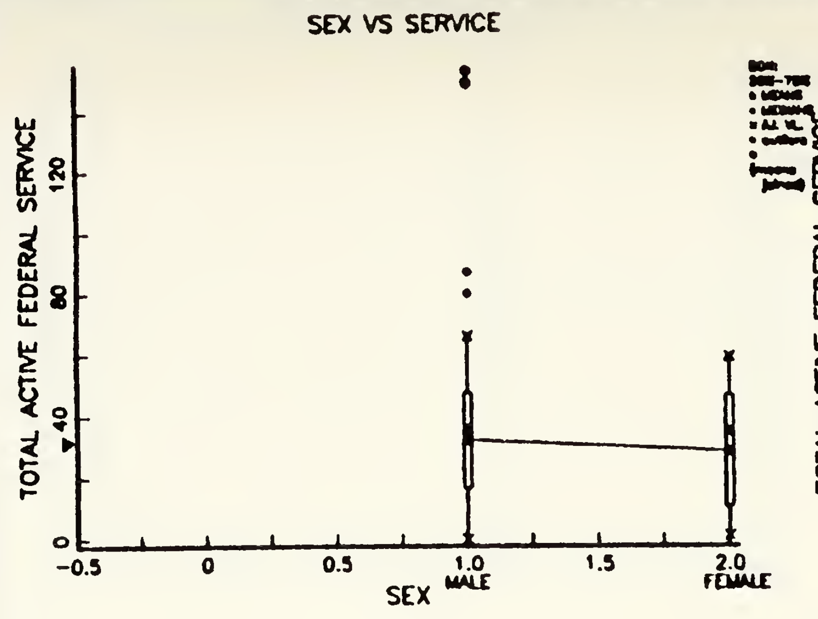
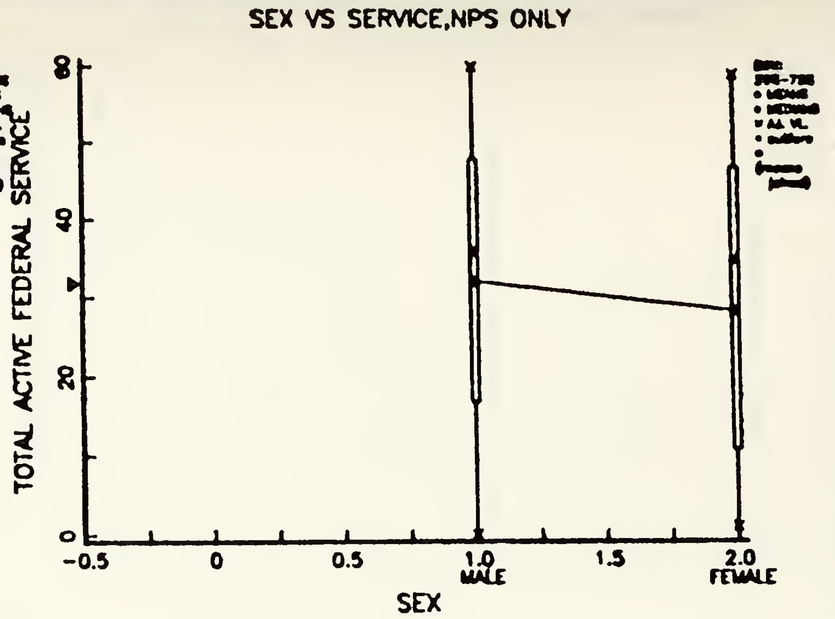


Figure H.24 Age vs. Service,IX



CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3]
Y :Y←A79MOD[;1]
SELECTION :ALL

CATEG.	NO.PTS	•/•−PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	31.975	17.838	18	38	48
1	2801	0.91001	32.32	17.768	17	38	48
2	277	0.089994	28.48	18.175	11	35	47



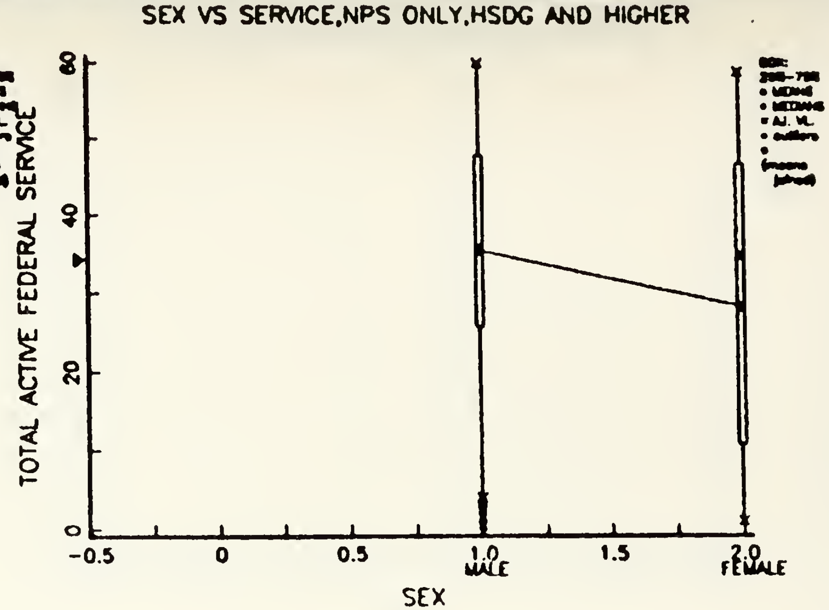
CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3
Y :Y←A79MOD[;1]
SELECTION :Y≤80

CATEG.	NO.PTS	•/•−PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	38	48
1	2796	0.90986	32.186	17.427	17	38	48
2	277	0.09014	28.48	18.175	11	35	47

Figure H.25 Sex vs. Service,I

CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3
Y :Y←A79MOD[;1]
SELECTION :Y≤60

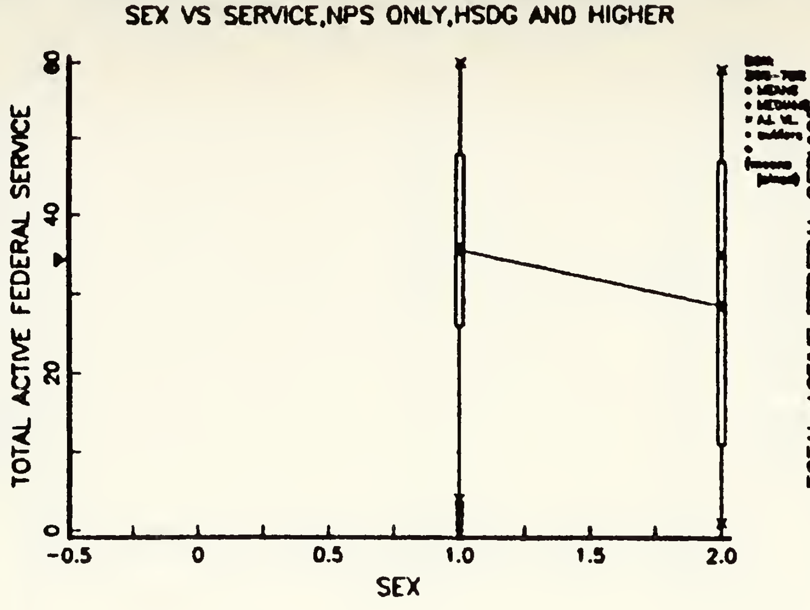
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	36	48
1	2796	0.90986	32.186	17.427	17	36	48
2	277	0.09014	28.48	18.175	11	35	47

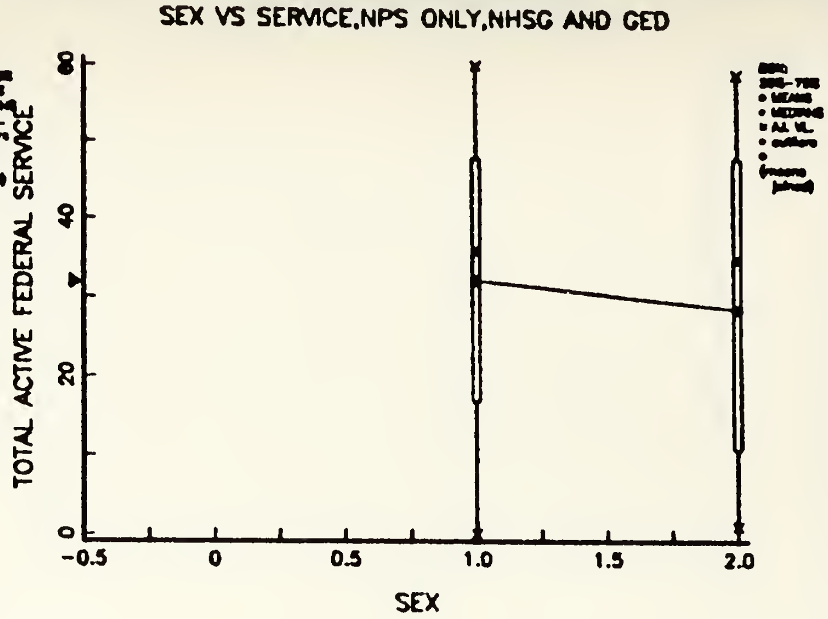


CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧((B79MOD[;2]>5)∧(B79MOD[;2]<13))

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.347	18.869	22	36	48
1	1337	0.8294	35.584	18.332	26	36	48
2	275	0.1706	28.444	18.135	11	35	47

Figure H.26 Sex vs. Service,II





CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧((B79MOD[;2]>5)∧(B79MOD[;2]<13))

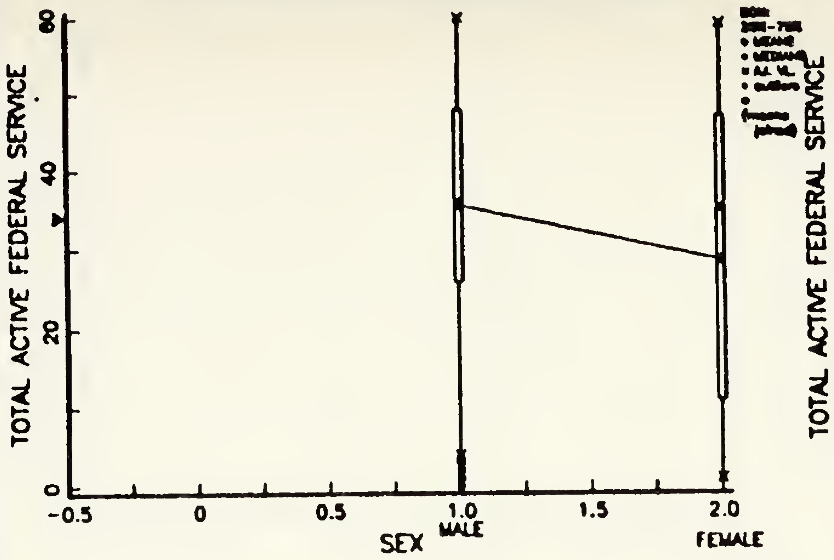
CATEG.	NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.387	18.869	22	36	48
1	1337	0.8294	35.586	18.332	26	36	48
2	275	0.1706	28.444	18.135	11	35	47

CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3]
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧((B79MOD[;2]≤8)∨(B79MOD[;2]≥13))

CATEG.	NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3004	1	31.922	17.531	16	36	48
1	2736	0.91079	32.235	17.435	17	36	48
2	268	0.089214	28.724	18.174	11	35	48

Figure H.27 Sex vs. Service,III

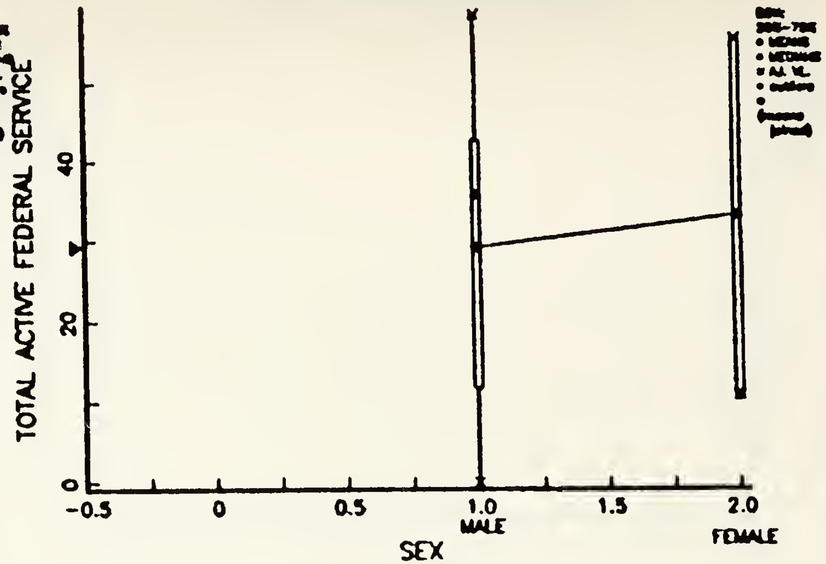
SEX VS SERVICE,NPS ONLY,HSDG AND HIGHER



CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧((B79MOD[;2]>5)∧(B79MOD[;2]<13))

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	38	48
1	1337	0.8294	35.586	16.332	28	38	48
2	275	0.1706	28.444	18.135	11	35	47

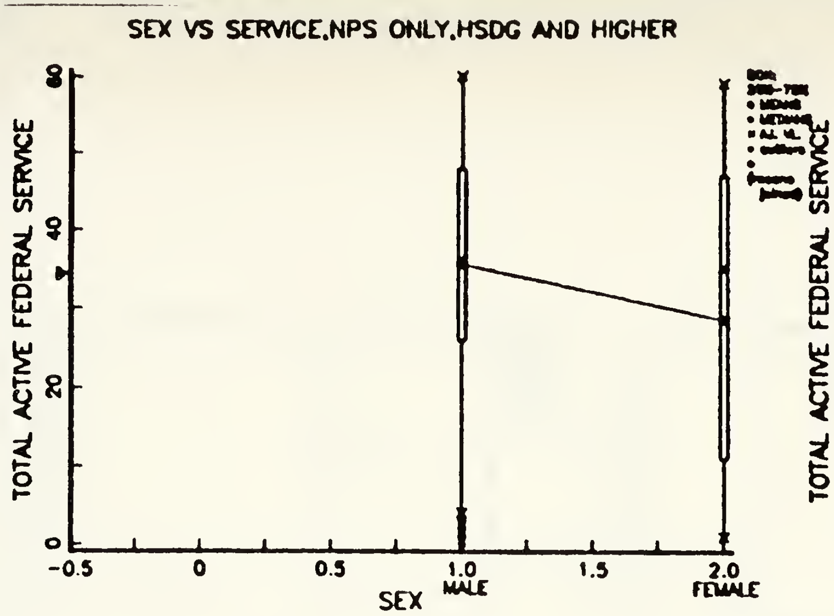
SEX VS SERVICE,NPS ONLY,NHSG



CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3]
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;2]<6)

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1265	1	29.372	17.748	12	36	43
1	1263	0.99842	29.365	17.730	12	36	43
2	2	0.001581	33.5	22.5	11	11	56

Figure H.28 Sex vs. Service,IV



CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)^((B79MOD[;2]>5)^(B79MOD[;2]<13))

CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	38	48
1	1337	0.8294	35.506	16.332	26	38	48
2	275	0.1706	28.444	18.135	11	35	47

CATEGORY VECTOR :(B79MOD[;3]≥1)+(B79MOD[;3]≥2)+(B79MOD[;3
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)^(B79MOD[;2]=13)

CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	196	1	27.173	18.159	10	30	40
1	196	1	27.173	18.159	10	30	40

Figure H.29 Sex vs. Service,V

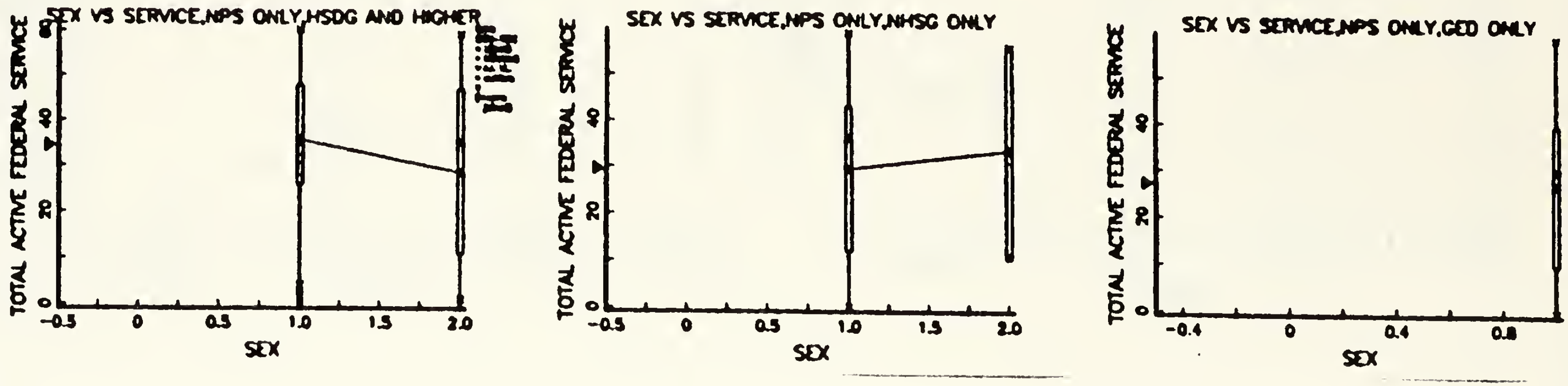
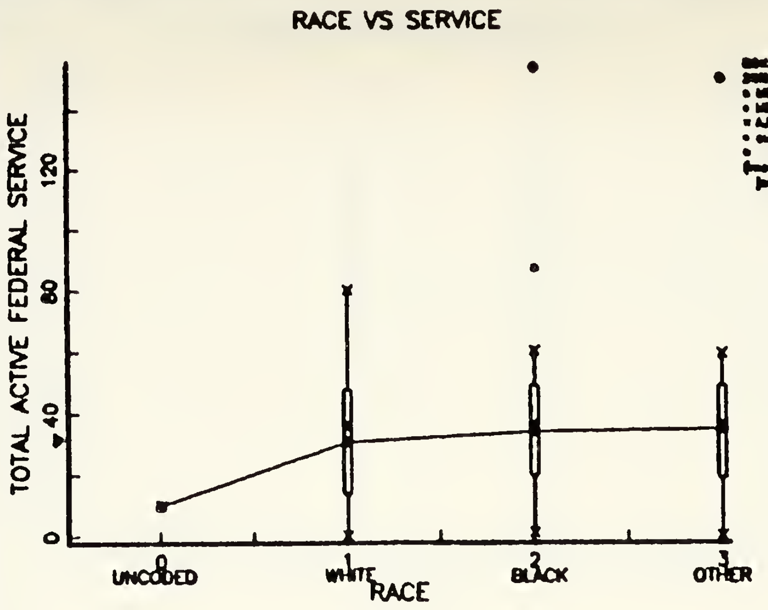
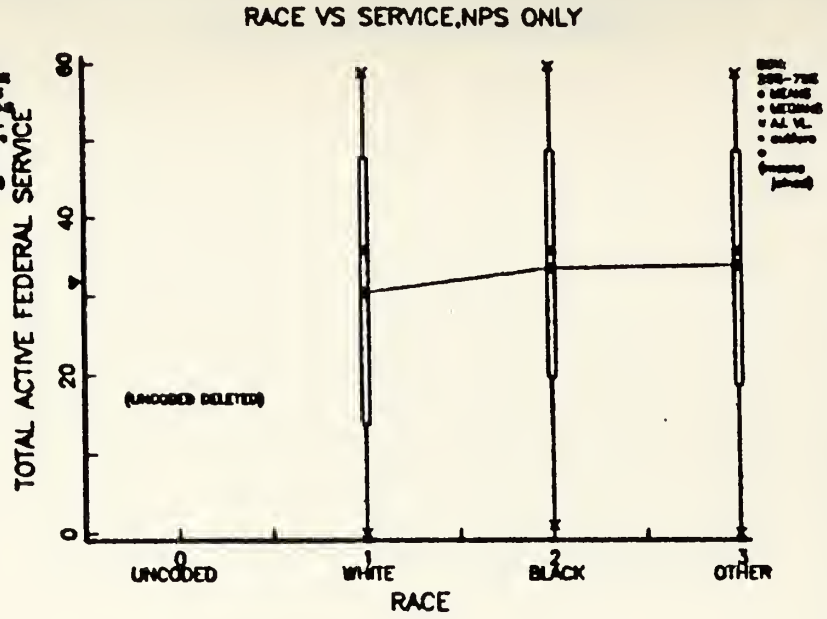


Figure H.30 Sex vs. Service,VI



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]≥3
Y :Y←A79MOD[;1]
SELECTION :ALL

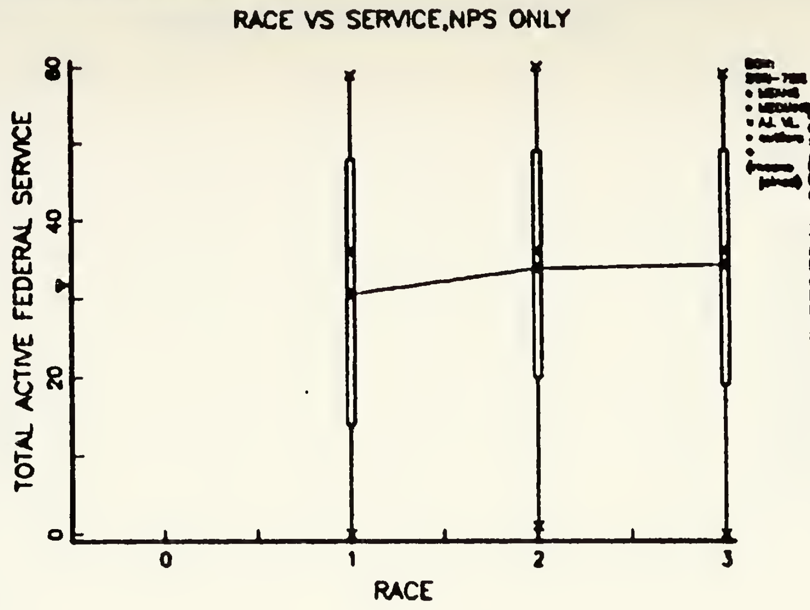
CATEG.		NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL		3078	1	31.975	17.838	18	38	48
	0	1	0.00032488	10	0	10	10	10
	1	1874	0.60884	30.823	17.557	14	38	48
	2	969	0.31481	33.965	17.978	20	38	49
	3	234	0.076023	34.854	18.395	19	38	49



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;4]≠0)

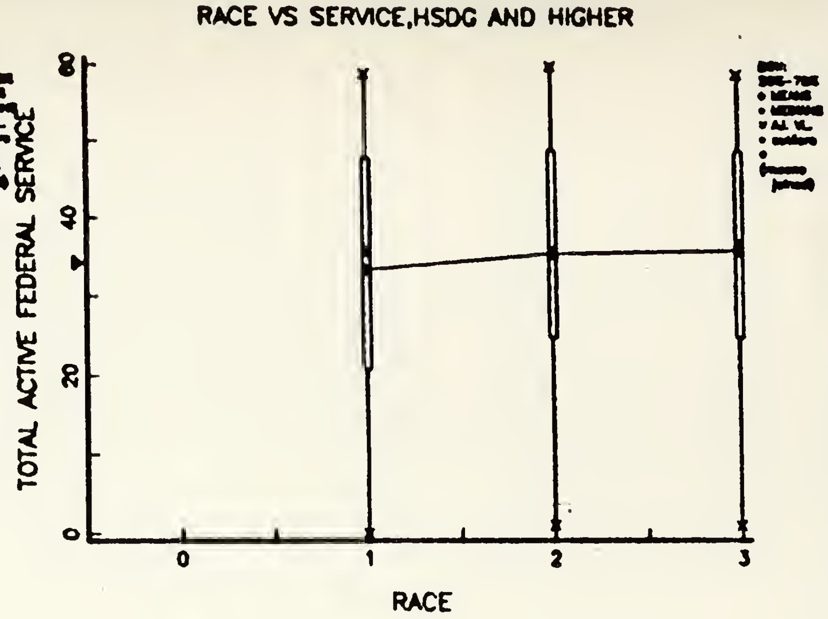
CATEG.		NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL		3072	1	31.858	17.528	18	38	48
	1	1872	0.60938	30.577	17.51	14	38	48
	2	967	0.31478	33.786	17.484	20	38	49
	3	·233	0.075846	34.158	18.807	18	38	49

Figure H.31 Race vs. Service,I



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;4]=0)

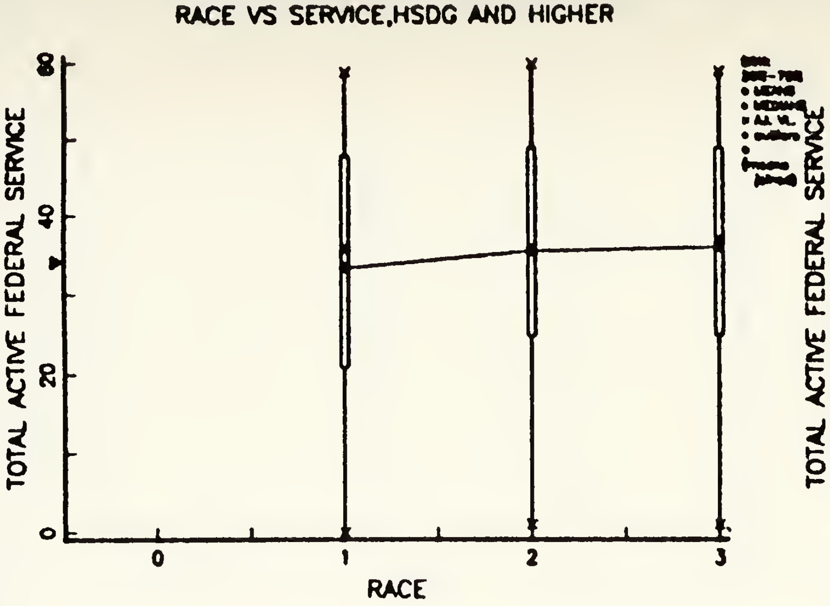
CATEG.		NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL		3072	1	31.859	17.526	18	36	48
	1	1872	0.60938	30.577	17.51	14	36	48
	2	967	0.31478	33.786	17.494	20	36	48
	3	233	0.075846	34.159	16.807	19	36	49



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :SELRC

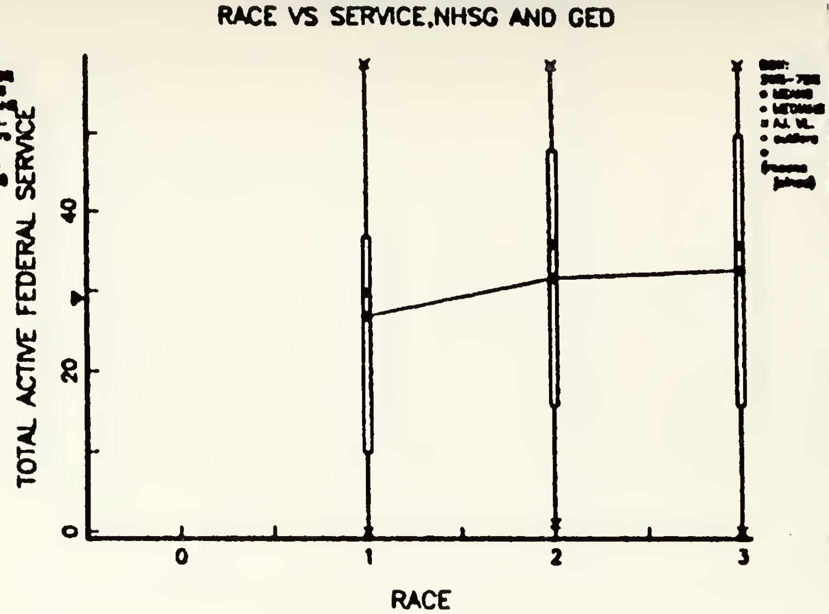
CATEG.		NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL		1612	1	34.367	16.889	22	36	48
	1	1024	0.63524	33.573	16.666	21	36	48
	2	493	0.30583	35.688	17.235	25	36	49
	3	95	0.058933	36.074	16.328	25	37	49

Figure H.32 Race vs. Service,II



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :SELRC

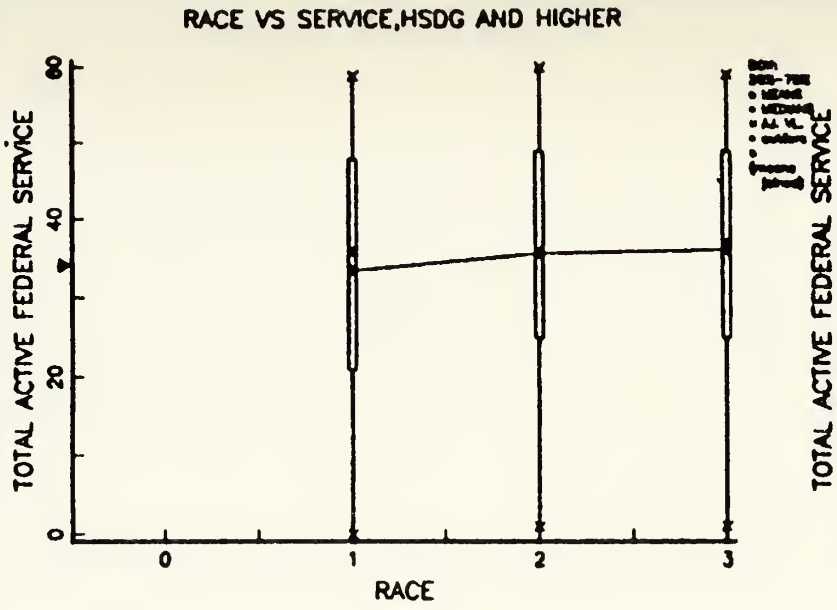
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	36	48
1	1024	0.63524	33.573	16.688	21	36	48
2	493	0.30583	35.688	17.235	25	36	49
3	95	0.058933	36.074	16.328	25	37	48



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :SLRC

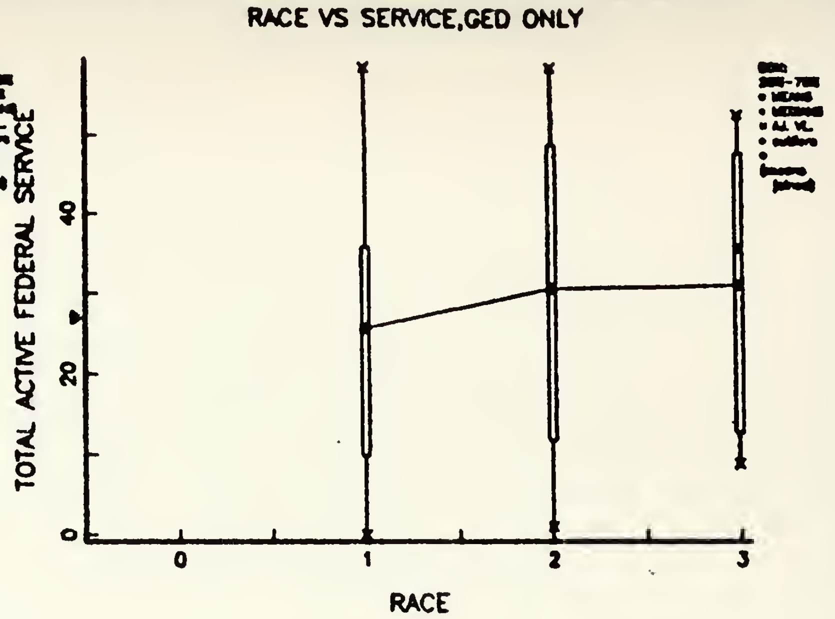
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1460	1	29.09	17.819	12	36	42
1	848	0.58082	26.96	17.795	10	30	37
2	474	0.32466	31.808	17.842	16	36	48
3	138	0.094521	32.841	17.004	16	36	50

Figure H.33 Race vs. Service,III



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :SELRC

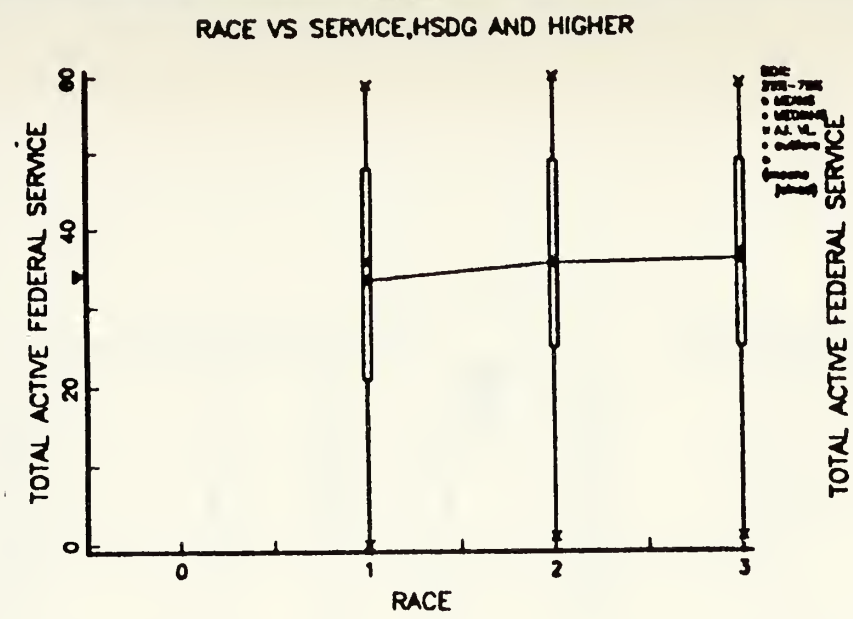
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	18.869	22	36	48
1	1024	0.63524	33.573	18.688	21	36	48
2	493	0.30583	35.688	17.235	25	36	49
3	95	0.058933	36.074	18.328	25	37	49



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :(Y≤80)∧(B79MOD[;4]≠0)∧(B79MOD[;2]=13)

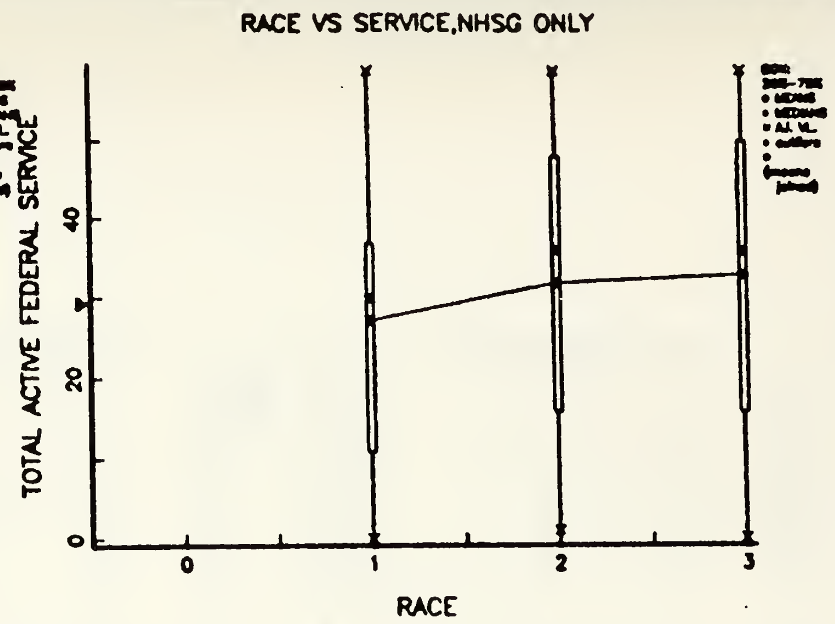
CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	196	1	27.173	18.159	10	30	40
1	144	0.73469	25.828	17.989	10	28	36
2	43	0.21939	30.814	18.414	12	31	40
3	9	0.045918	31.333	18.553	13	36	48

Figure H.34 Race vs. Service,IV



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4]
Y :Y←A79MOD[;1]
SELECTION :SELRC

CATEG.	NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.347	16.849	22	36	48
1	1024	0.63524	33.573	16.688	21	36	48
2	493	0.30583	35.888	17.235	25	38	49
3	95	0.058933	36.074	16.328	25	37	49



CATEGORY VECTOR :(B79MOD[;4]≥1)+(B79MOD[;4]≥2)+(B79MOD[;4
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(B79MOD[;4]≠0)∧(B79MOD[;2]≤5)

CATEG.	NO.PTS	∘/∘-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1264	1	29.367	17.747	12	38	43
1	704	0.55696	27.182	17.748	11	30	37
2	431	0.34098	31.907	17.449	16	36	48
3	129	0.10206	32.946	17.03	16	36	50

Figure H.35 Race vs. Service,V

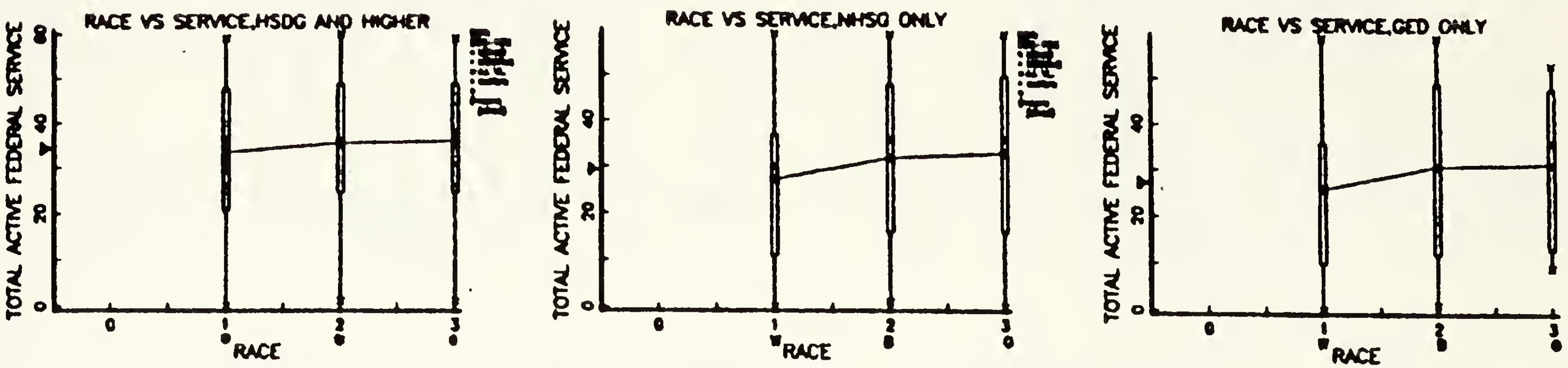
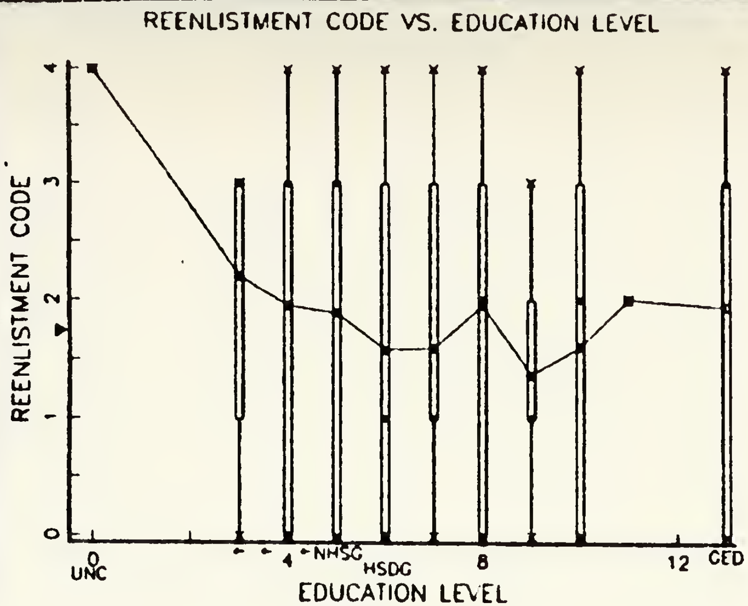
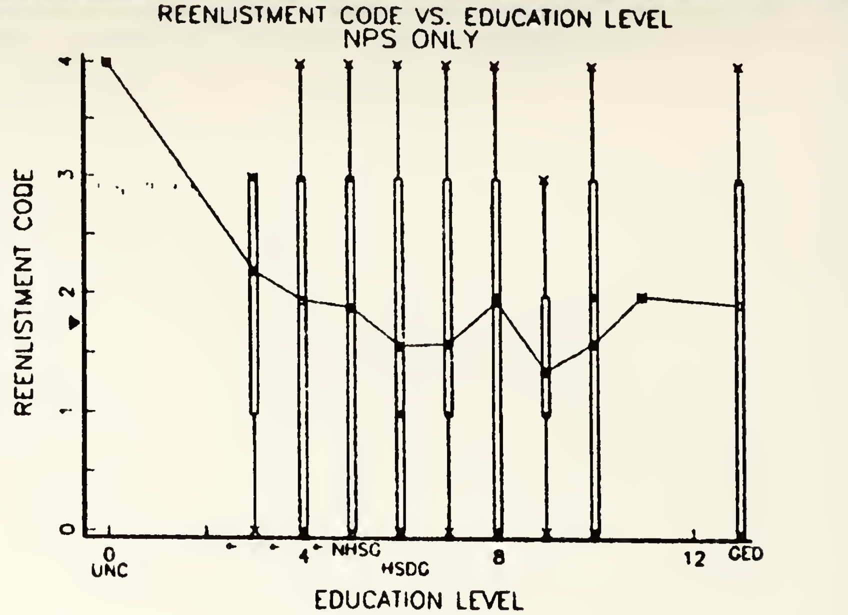


Figure H.36 Race vs. Service,VI





CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;3]
SELECTION :ALL

X LABEL :EDUC
Y LABEL :REEN
NO. OF ELEMENTS :3078

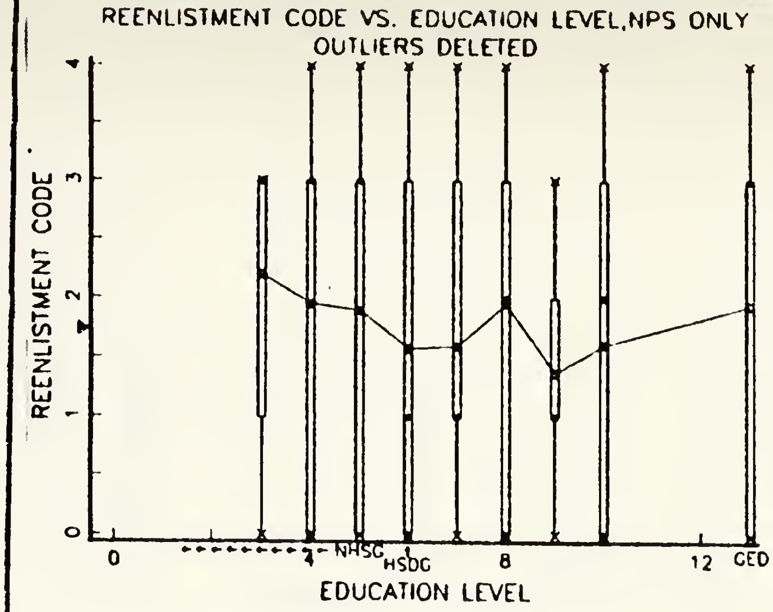
CATEG.	NO.PTS	●/●-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	1.7453	1.3845	0	2	3
0	1	0.00032489	4	0	4	4	4
3	10	0.0032489	2.2	1.249	1	3	3
4	752	0.24431	1.9495	1.4077	0	3	3
5	504	0.16374	1.8889	1.4294	0	3	3
6	1546	0.50227	1.5737	1.3371	0	1	3
7	22	0.0071475	1.5909	1.1931	1	1	3
8	20	0.0064977	1.95	1.4309	0	2	3
9	11	0.0035737	1.3636	0.8814	1	1	2
10	15	0.0048733	1.6	1.2543	0	2	3
11	1	0.00032489	2	0	2	2	2
13	196	0.063678	1.9388	1.4129	0	3	3

CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;3]
SELECTION :(A79MOD[;1])≤60

X LABEL :E
Y LABEL :R
NO. OF ELEMENTS :3

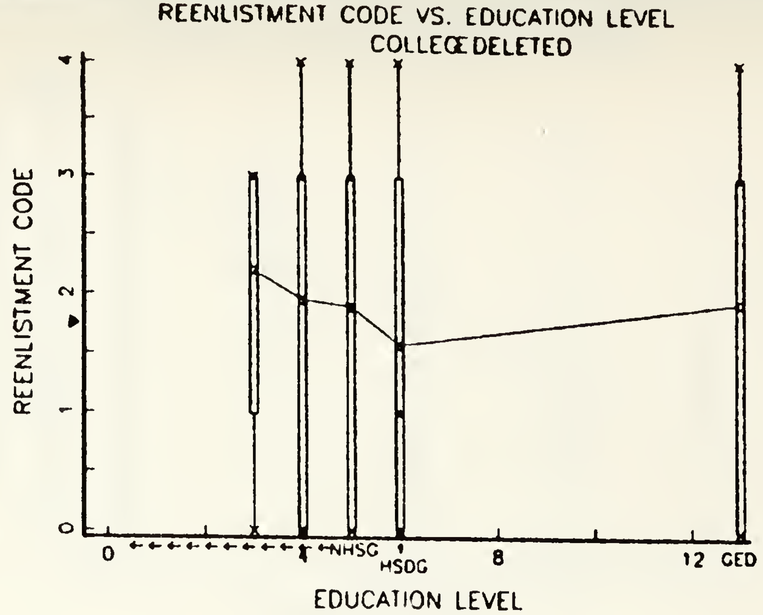
CATEG.	NO.PTS	●/●-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	1.7452	1.3844	0	2	3
0	1	0.00032541	4	0	4	4	4
3	10	0.0032541	2.2	1.249	1	3	3
4	751	0.24439	1.9521	1.4068	0	3	3
5	503	0.16368	1.8926	1.4283	0	3	3
6	1543	0.50212	1.571	1.3369	0	1	3
7	22	0.0071591	1.5909	1.1931	1	1	3
8	20	0.0065083	1.95	1.4309	0	2	3
9	11	0.0035796	1.3636	0.8814	1	1	2
10	15	0.0048812	1.6	1.2543	0	2	3
11	1	0.00032541	2	0	2	2	2
13	188	0.063781	1.9388	1.4129	0	3	3

Figure H.37 Reenlistment Code vs. Education Level,I



CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;3]
SELECTION :((A79MOD[;1])≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;

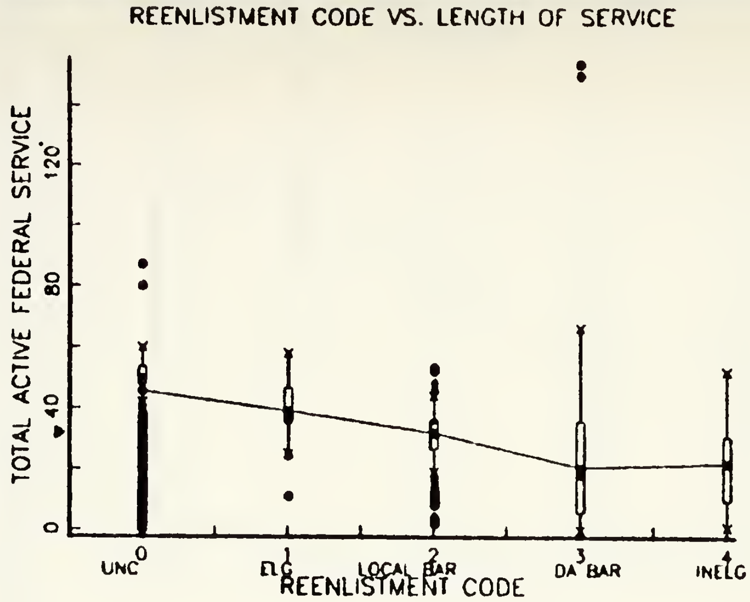
CATEG.	NO.PTS	ø/ø-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3071	1	1.7444	1.3842	0	2	3
3	10	0.0032563	2.2	1.249	1	3	3
4	751	0.24455	1.9521	1.4068	0	3	3
5	503	0.16379	1.8926	1.4283	0	3	3
6	1543	0.50244	1.571	1.3369	0	1	3
7	22	0.0071638	1.5909	1.1931	1	1	3
8	20	0.0065125	1.95	1.4309	0	2	3
9	11	0.0035819	1.3636	0.8814	1	1	2
10	15	0.0048844	1.6	1.2543	0	2	3
13	196	0.063823	1.9388	1.4129	0	3	3



CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;3]
SELECTION :((A79MOD[;1])≤60)∧(B79MOD[;2]≠0)∧(B79MOD[

CATEG.	NO.PTS	ø/ø-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3003	1	1.7463	1.3869	0	2	3
3	10	0.00333	2.2	1.249	1	3	3
4	751	0.25008	1.9521	1.4068	0	3	3
5	503	0.1675	1.8926	1.4283	0	3	3
6	1543	0.51382	1.571	1.3369	0	1	3
13	196	0.065268	1.9388	1.4129	0	3	3

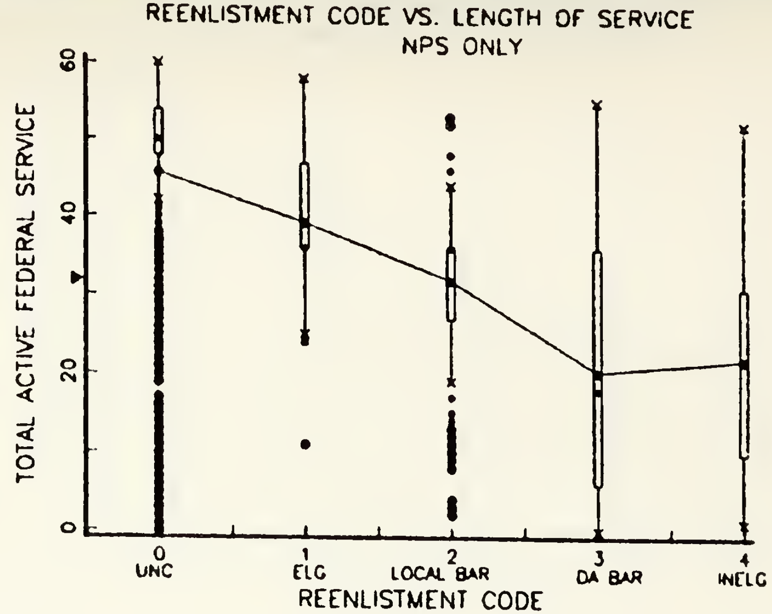
Figure H.38 Reenlistment Code vs. Education Level,II



CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :ALL

X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :30

CATEG.	NO.PTS	I•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	31.975	17.838	16	36	48
0	926	0.30084	45.677	14.225	48	50	54
1	523	0.16992	39.143	5.7839	36	36	47
2	179	0.058155	31.816	9 5937	27	36	36
3	1309	0.42528	20.523	16 261	6	18	36
4	141	0.045809	21.908	13.395	10	22	31

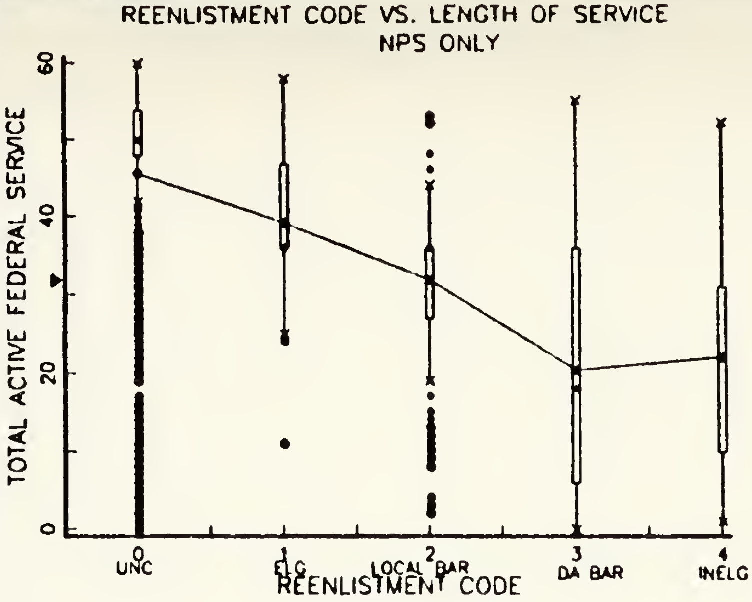


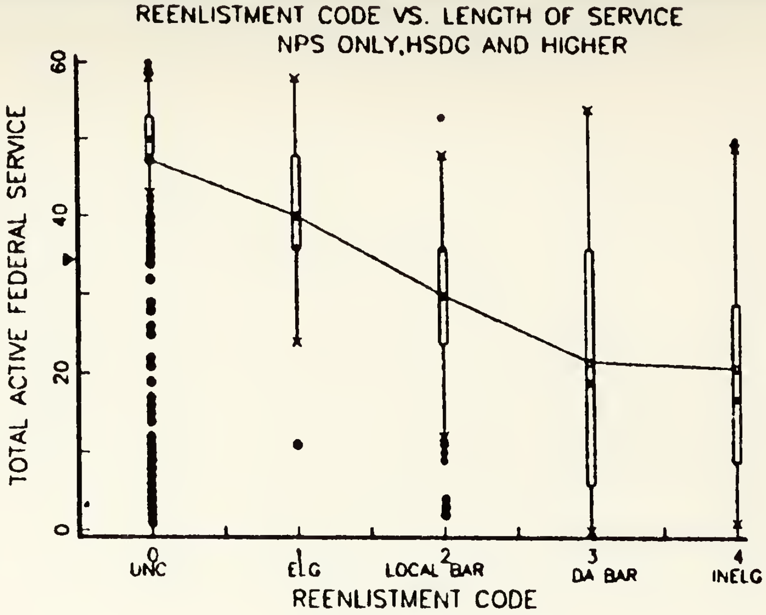
CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :Y≤60

X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :30

CATEG.	NO.PTS	I•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31 852	17.528	16	36	48
0	924	0.30068	45.595	14.13	48	50	54
1	523	0.17019	39.143	5.7839	36	36	47
2	179	0.058249	31.816	9.5937	27	36	36
3	1306	0 42499	20.287	15.391	6	18	36
4	141	0.045884	21.908	13.395	10	22	31

Figure H.39 Reenlistment Code vs. Service,I





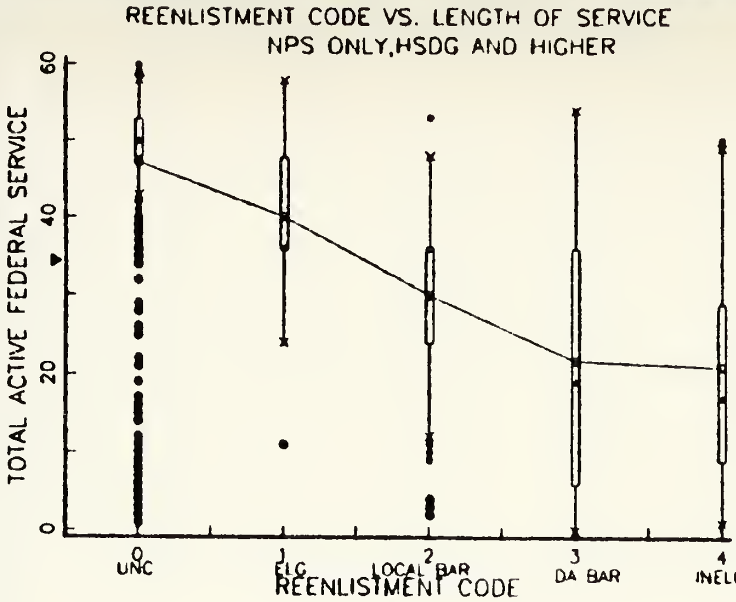
CATEGORY VECTOR :RCAT
Y :Y-A79MOD[;1]
SELECTION :Y≤60
X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :30

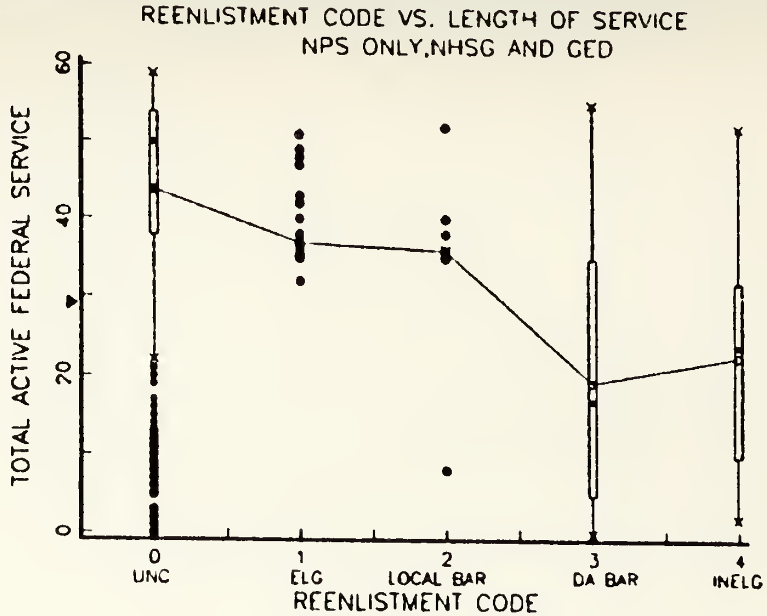
CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	31.852	17.528	18	36	48
0	924	0.30068	45.595	14.13	48	50	54
1	523	0.17019	39.143	5.7839	36	36	47
2	179	0.058249	31.818	9.5937	27	36	36
3	1306	0.42499	20.287	15.391	6	18	36
4	141	0.045884	21.908	13.395	10	22	31

CATEGORY VECTOR :RCAT
Y :Y-A79MOD[;1]
SELECTION :HSG
X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :16

CATEG.	NO.PTS	•/•-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	36	48
0	498	0.30893	47.209	11.9	48	50	53
1	369	0.22891	40.095	6.3808	36	36	48
2	123	0.076303	29.951	10.672	24	36	36
3	564	0.34988	21.628	16.285	6	19	36
4	58	0.03598	20.914	13.718	9	17	29

Figure H.40 Reenlistment Code vs. Service,II





CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :HSG

X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :16

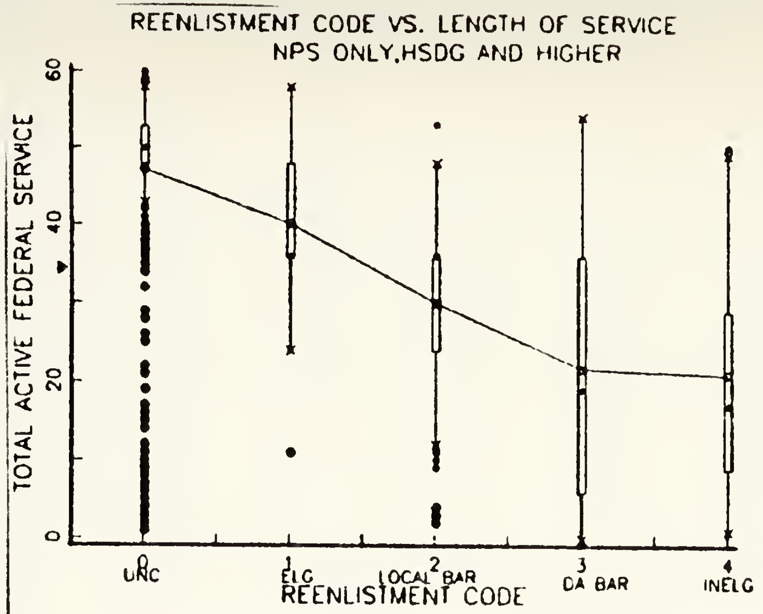
CATEG.	NO.PTS	●/●-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1812	1	34.367	16.869	22	36	48
0	498	0.30893	47.209	11.9	48	50	53
1	369	0.22891	40.095	8.3808	36	36	48
2	123	0.076303	29.951	10.672	24	36	36
3	564	0.34988	21.628	16.285	6	19	36
4	58	0.03598	20.914	13.718	9	17	29

CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :NHSGED

X LABEL :R
Y LABEL :T
NO. OF ELEMENTS :1

CATEG	NO.PTS	●/●-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1461	1	29.077	17.82	12	36	42
0	426	0.29158	43.709	16.153	38	50	54
1	154	0.10541	36.864	2.9477	36	36	36
2	56	0.03833	35.911	4.4291	36	36	36
3	742	0.50787	19.268	14.594	5	17	35
4	83	0.05681	22.602	13.12	10	24	32

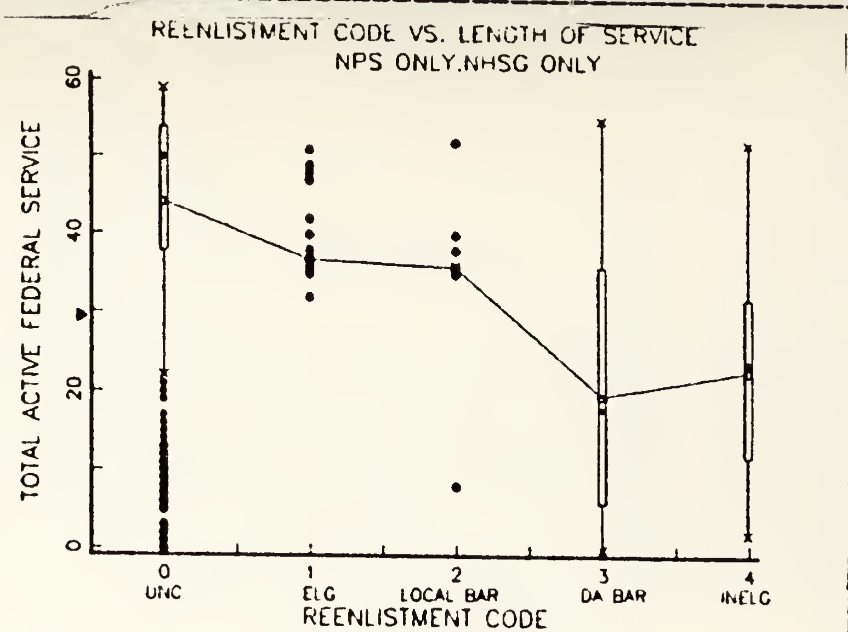
Figure H.41 Reenlistment Code vs. Service,III



CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :HSG

X LABEL :RE
Y LABEL :TO
NO OF ELEMENTS :16

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	36	48
0	498	0.30893	47.209	11.9	48	50	53
1	369	0.22891	40.095	6.3808	36	36	48
2	123	0.076303	29.951	10.672	24	38	36
3	564	0.34988	21.628	16.285	8	19	36
4	58	0.03598	20.914	13.718	9	17	29

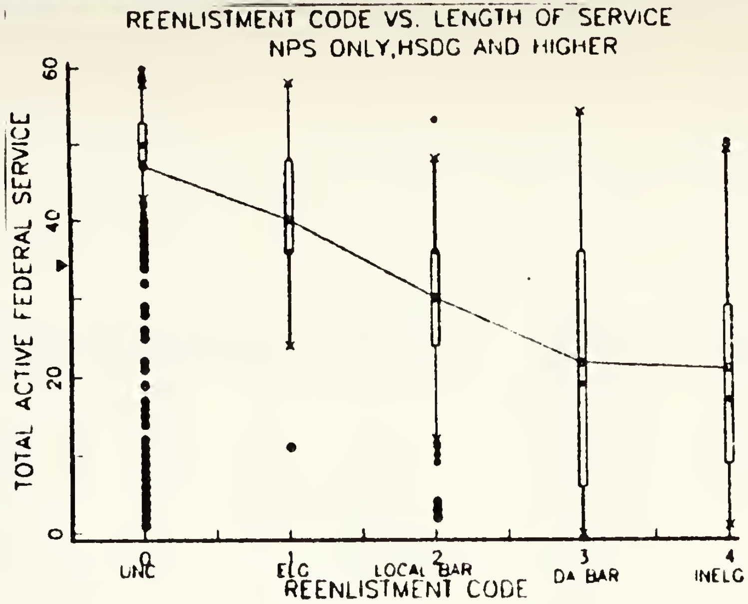


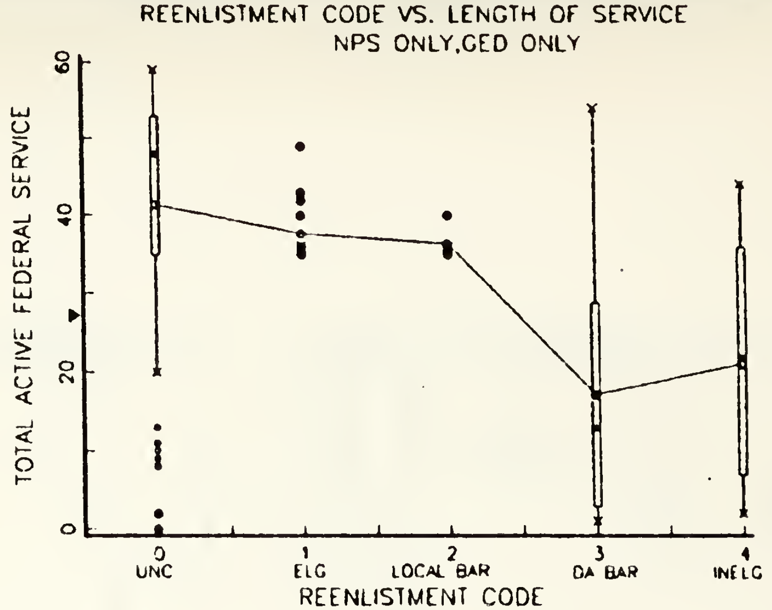
CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :NHSG

X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :12

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1265	1	29.372	17.748	12	36	43
0	368	0.29091	44.087	16.014	38	50	54
1	136	0.10751	35.765	2.8472	36	36	36
2	48	0.037945	35.833	4.7449	36	36	36
3	640	0.50593	19.603	14.603	6	18	36
4	73	0.057708	22.808	12.878	12	24	32

Figure H.42 Reenlistment Code vs. Service,IV





CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :HSG

X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :18

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1612	1	34.367	16.869	22	36	48
0	498	0.30893	47.209	11.9	48	50	53
1	369	0.22891	40.095	6.3808	36	36	48
2	123	0.076303	29.951	10.672	24	36	36
3	564	0.34988	21.628	16.285	6	19	36
4	58	0.03598	20.914	13.718	9	17	29

CATEGORY VECTOR :RCAT
Y :Y←A79MOD[;1]
SELECTION :GED

X LABEL :RE
Y LABEL :TO
NO. OF ELEMENTS :19

CATEG.	NO.PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	196	1	27.173	18.159	10	30	40
0	58	0.29592	41.31	16.813	35	48	53
1	18	0.091837	37.611	3.5299	36	36	36
2	8	0.040816	36.375	1.4087	36	36	36
3	102	0.52041	17.167	14.358	3	13	29
4	10	0.05102	21.1	14.68	7	22	36

Figure H.43 Reenlistment Code vs. Service,V

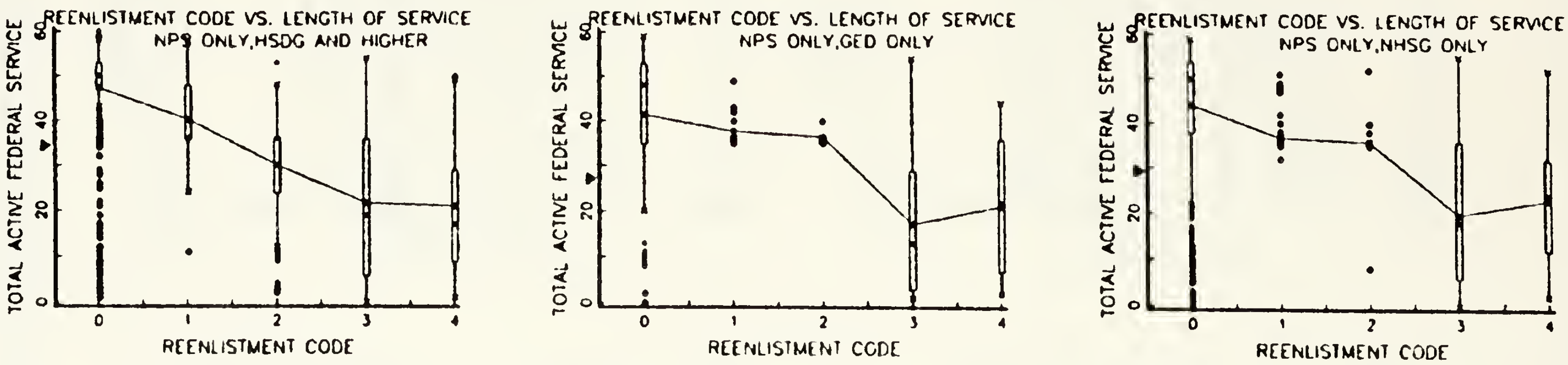
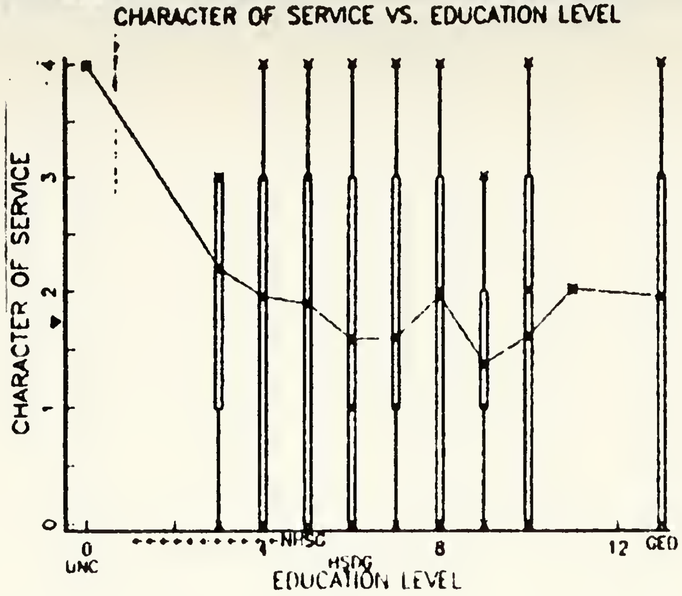
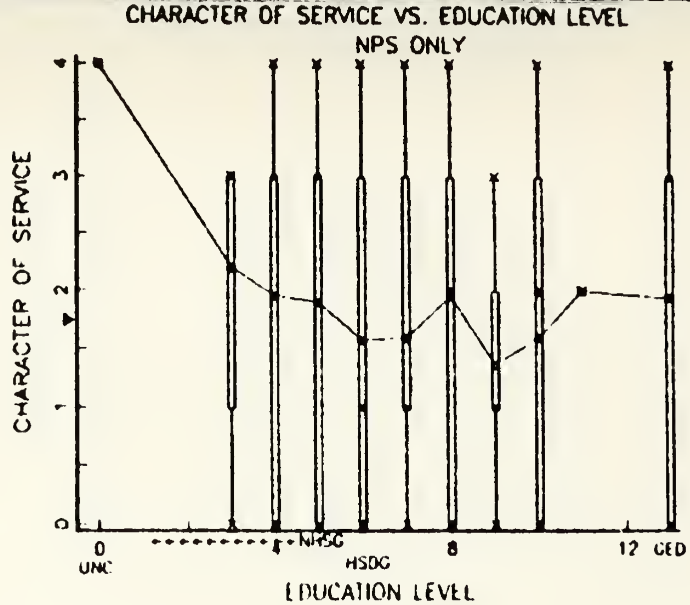


Figure H.44 Reenlistment Code vs. Service,VI





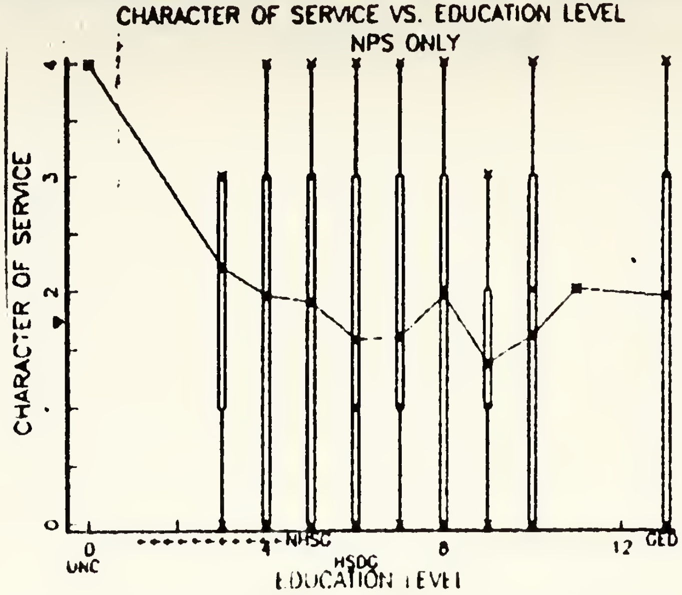
CATEGORY VECTOR :EDCAT X LABEL :EDUC
Y :Y←A79MOD[;3] Y LABEL :CHAR
SELECTION :ALL NO OF ELEMENTS :3078

CATEG	NO PTS	l+/+-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3078	1	1.7453	1.3845	0	2	3
0	1	0.00032489	4	0	4	4	4
3	10	0.0032489	2.2	1.249	1	3	3
4	752	0.24431	1.9495	1.4077	0	3	3
5	504	0.16374	1.8889	1.4294	0	3	3
6	1546	0.50227	1.5737	1.3371	0	1	3
7	22	0.0071475	1.5909	1.1831	1	1	3
8	20	0.0064977	1.95	1.4309	0	2	3
9	11	0.0035737	1.3636	0.8814	1	1	2
10	15	0.0048733	1.6	1.2543	0	2	3
11	1	0.00032489	2	0	2	2	2
13	196	0.063678	1.9388	1.4129	0	3	3

CATEGORY VECTOR :EDCAT X LABEL :E
Y :Y←A79MOD[;3] Y LABEL :C
SELECTION :(A79MOD[;1])≠80 NO OF ELEMENTS :3

CATEG	NO PTS	l+/+-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	3073	1	1.7452	1.3844	0	2	3
0	1	0.00032541	4	0	4	4	4
3	10	0.0032541	2.2	1.249	1	3	3
4	751	0.24439	1.9521	1.4068	0	3	3
5	503	0.16368	1.8926	1.4283	0	3	3
6	1543	0.50212	1.571	1.3369	0	1	3
7	22	0.0071591	1.5909	1.1831	1	1	3
8	20	0.0065083	1.95	1.4309	0	2	3
9	11	0.0035796	1.3636	0.8814	1	1	2
10	15	0.0048812	1.6	1.2543	0	2	3
11	1	0.00032541	2	0	2	2	2
13	196	0.063781	1.9388	1.4129	0	3	3

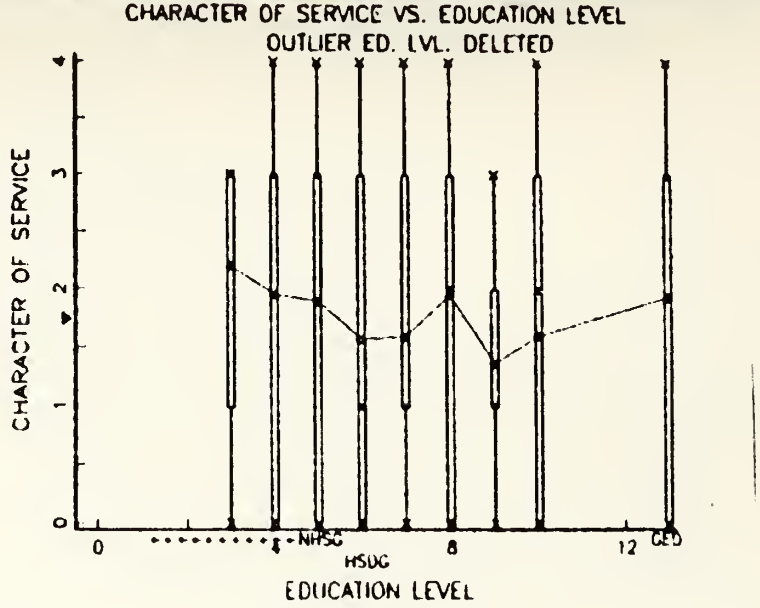
Figure B.45 Character of Service vs. Education Level,I



CATEGORY VECTOR EDCAT
Y :Y←A79MOD[,3]
SELECTION :(A79MOD[,1])≤60

X LABEL :E
Y LABEL :C
NO OF ELEMENTS :3

CATEG	NO PTS	o/o-PTS	Y MEAN	Y DVN	.25	50	.75
ALL	3073	1	1 7452	1 3844	0	2	3
0	1	0 00032541	4	0	4	4	4
3	10	0 0032541	2 2	1 249	1	3	3
4	751	0 24439	1.9521	1 4068	0	3	3
5	503	0 16368	1 8926	1 4283	0	3	3
6	1543	0 50212	1 571	1.3368	0	1	3
7	22	0 0071591	1 5909	1 1931	1	1	3
8	20	0 0065083	1 95	1 4309	0	2	3
9	11	0 0035798	1 3636	0 8814	1	1	2
10	15	0 0048812	1 6	1 2543	0	2	3
11	1	0 00032541	2	0	2	2	2
13	196	0 063781	1 9388	1 4179	0	3	3

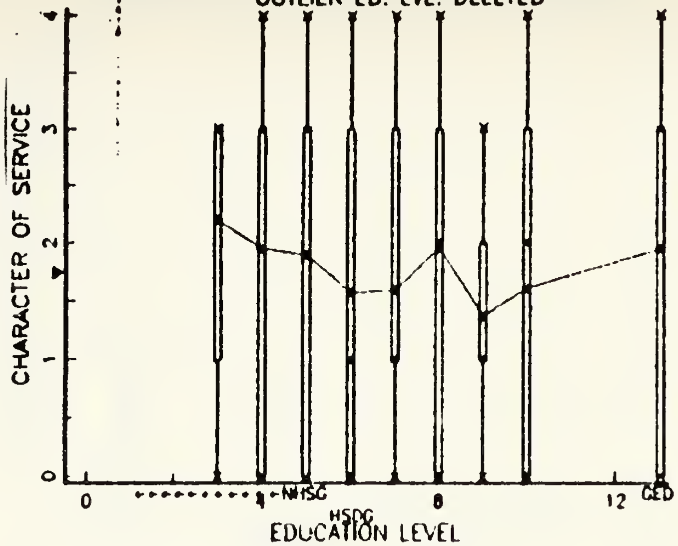


CATEGORY VECTOR .EDCAT
Y Y←A79MOD[,3]
SELECTION .(A79MOD[:1]≤60)∧(B79MOD[,2]=0)∧(B79MOD[:2]

CATEG	NO PTS	o/o-PTS	Y MEAN	Y DVN	25	50	75
ALL	3071	1	1 7444	1 3842	0	2	3
3	10	0 0032563	2 2	1 249	1	3	3
4	751	0 24455	1 9521	1 4068	0	3	3
5	503	0 16379	1 8926	1 4283	0	3	3
6	1543	0 50244	1 571	1 3368	0	1	3
7	22	0 0071638	1 5909	1 1931	1	1	3
8	20	0 0065125	1 95	1 4309	0	2	3
9	11	0 0035818	1 3636	0 8814	1	1	2
10	15	0 0048844	1 6	1 2543	0	2	3
13	196	0 063823	1 9388	1 4129	0	3	3

Figure B.46 Character of Service vs. Education Level,II

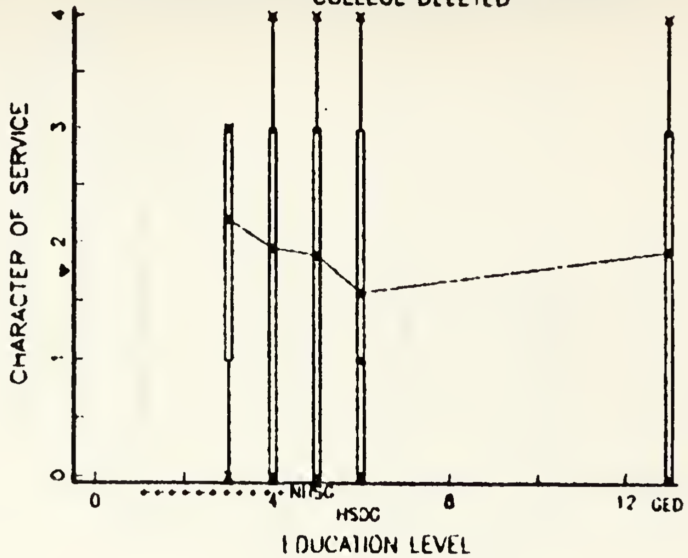
CHARACTER OF SERVICE VS. EDUCATION LEVEL
OUTLIER ED. LVL. DELETED



CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;3]
SELECTION :(A79MOD[;1]≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;2]

CATEG.	NO PTS	%-PTS	Y MEAN	Y DVN	.25	.50	.75
ALL	3071	1	1.7444	1.3842	0	2	3
3	10	0.0032563	2.2	1.249	1	3	3
4	751	0.24455	1.9521	1.4068	0	3	3
5	503	0.16379	1.8926	1.4283	0	3	3
6	1543	0.50244	1.571	1.3369	0	1	3
7	22	0.0071638	1.5909	1.1831	1	1	3
8	20	0.0065125	1.95	1.4309	0	2	3
9	11	0.0035819	1.3636	0.6814	1	1	2
10	15	0.0048844	1.6	1.2543	0	2	3
13	196	0.063823	1.9388	1.4129	0	3	3

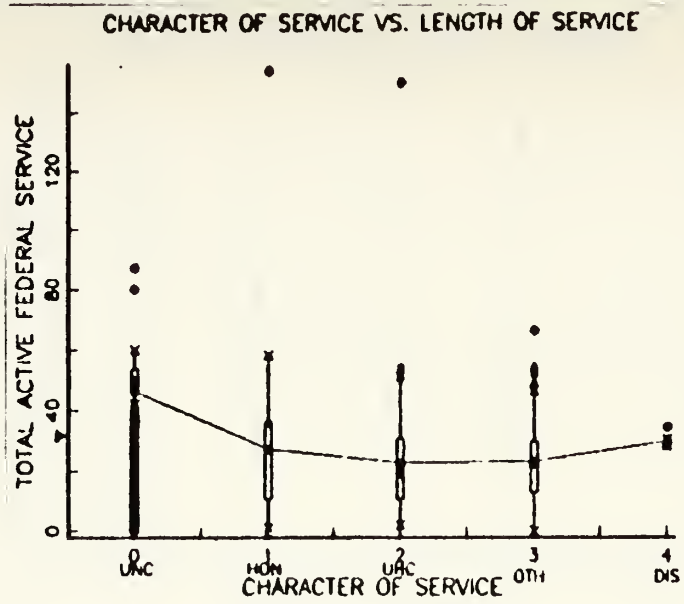
CHARACTER OF SERVICE VS. EDUCATION LEVEL
COLLEGE DELETED

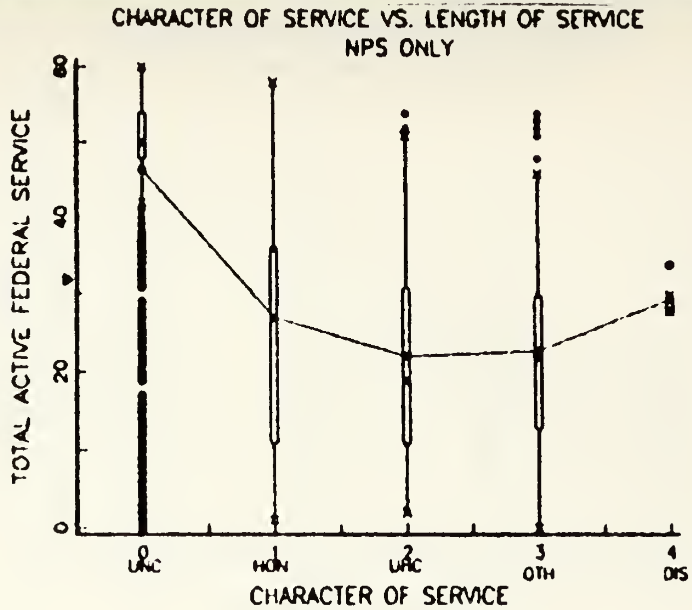


CATEGORY VECTOR :EDCAT
Y :Y←A79MOD[;3]
SELECTION :(A79MOD[;1]≤60)∧(B79MOD[;2]≠0)∧(B79MOD[;2

CATEG	NO.PTS	%-PTS	Y MEAN	Y DVN	.25	.50	.75
ALL	3003	1	1.7403	1.3869	0	2	3
3	10	0.00333	2.2	1.249	1	3	3
4	751	0.25008	1.9521	1.4068	0	3	3
5	503	0.1675	1.8926	1.4283	0	3	3
6	1543	0.51382	1.571	1.3369	0	1	3
13	196	0.065268	1.9388	1.4128	0	3	3

Figure H.47 Character of Service vs. Education Level,III





CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :ALL

X LABEL :CHA
Y LABEL :TOT
NO OF ELEMENTS :307

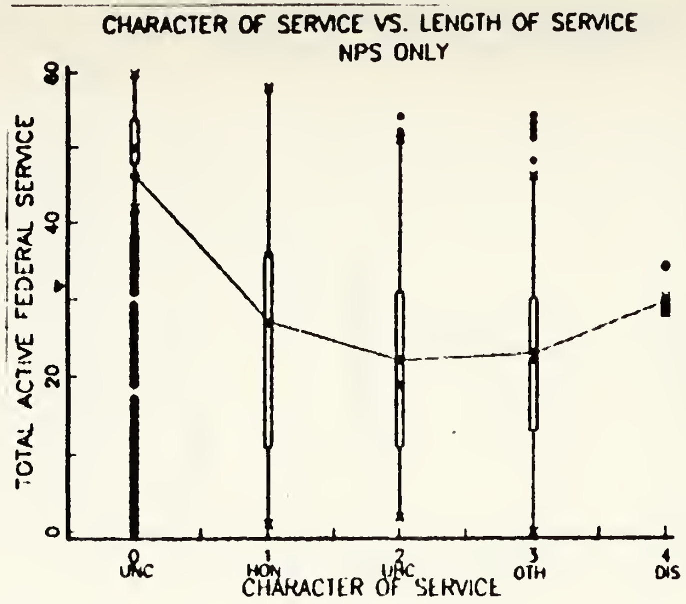
CATEG.	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	25	50	.75
ALL.	3078	1	31 875	17 838	16	36	48
0	890	0 28915	48.43	13 97	48	50	54
1	1684	0 54711	27 069	18 199	11	36	38
2	253	0 082196	22 548	15 372	11	19	31
3	245	0 079597	22 971	11 887	13	22	30
4	6	0 0019493	29 5	2 1409	28	28	30

CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :Y≤60

X LABEL CHA
Y LABEL :TOT
NO. OF ELEMENTS :307

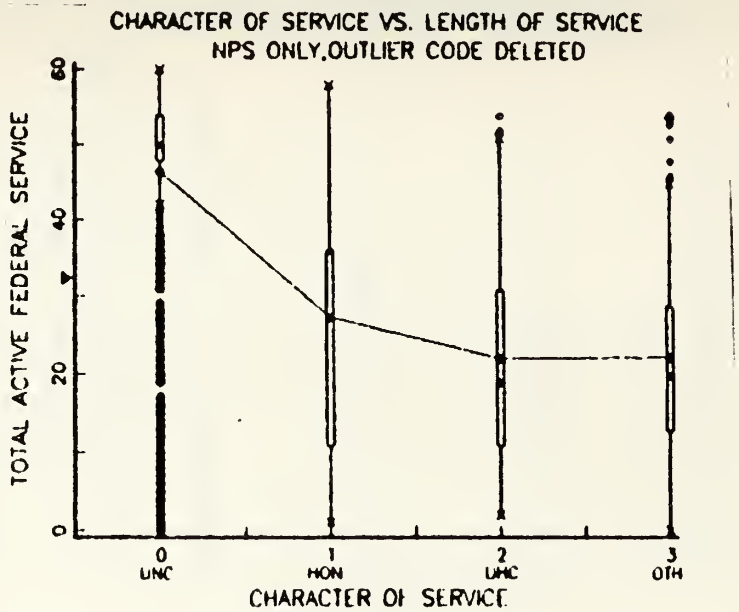
CATEG	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	25	.50	75
ALL	3073	1	31 852	17 528	18	36	48
0	888	0 28897	46 347	13 874	48	50	54
1	1683	0.54787	26 894	15 905	11	38	38
2	252	0 082005	22.044	13 134	11	19	31
3	244	0 079401	22 795	11 69	13	22	30
4	8	0 0018525	29 5	2 1409	28	28	30

Figure H. 48 Character of Service vs. Service, I



CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :Y≤60
X LABEL :CHA
Y LABEL :TOT
NO. OF ELEMENTS :307

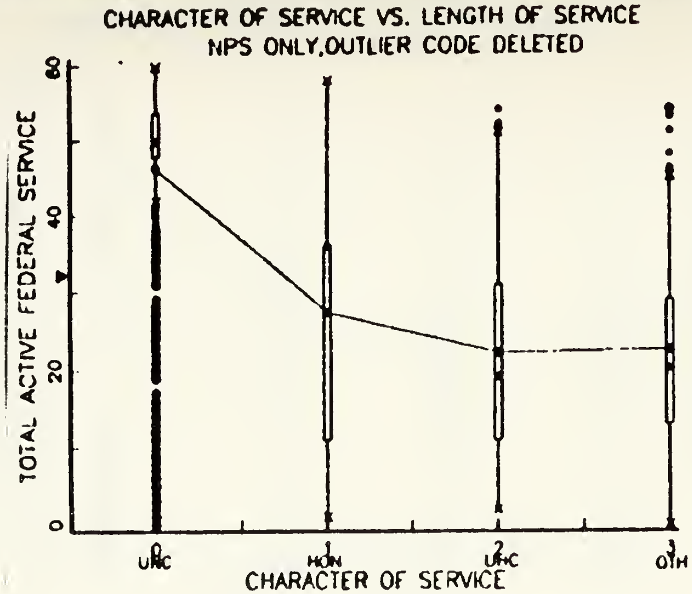
CATEG.	NO PTS	+/÷PTS	Y-MEAN	Y DVN	25	50	75
ALL	3073	1	31 852	17 528	16	36	48
0	888	0 28897	46 347	13 874	48	50	54
1	1683	0 54767	26 944	15 905	11	36	36
2	252	0 082005	22 044	13 134	11	19	31
3	244	0.079401	22 795	11 69	13	22	30
4	6	0.0018525	29 5	2 1409	28	28	30

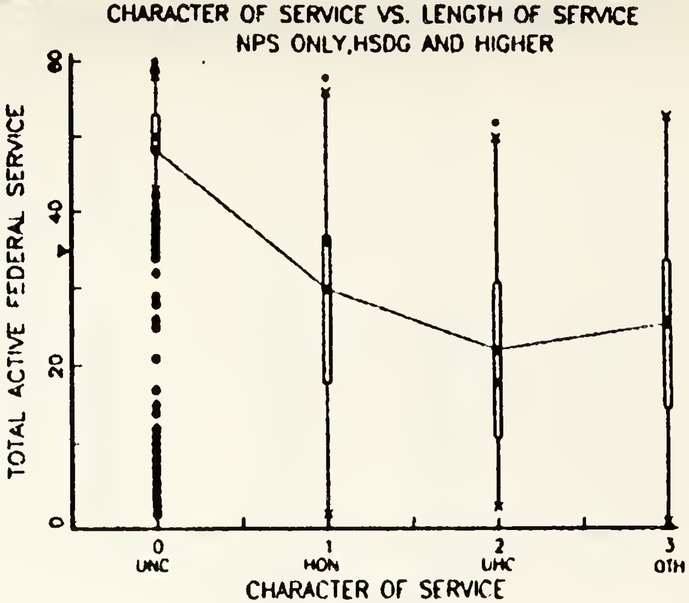


CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]≠4)
X LABEL
Y LABEL
NO. OF ELE

CATEG.	NO PTS	+/÷PTS	Y-MEAN	Y-DVN	25	50	75
ALL	2932	1	32 33	17 581	17	36	48
0	888	0 30286	46 347	13 874	48	50	54
1	1616	0.55116	27 317	15 862	11	36	36
2	248	0 084584	22 088	13 134	11	18	31
3	180	0 061392	22 322	11.707	13	20	29

Figure H.49 Character of Service vs. Service,II





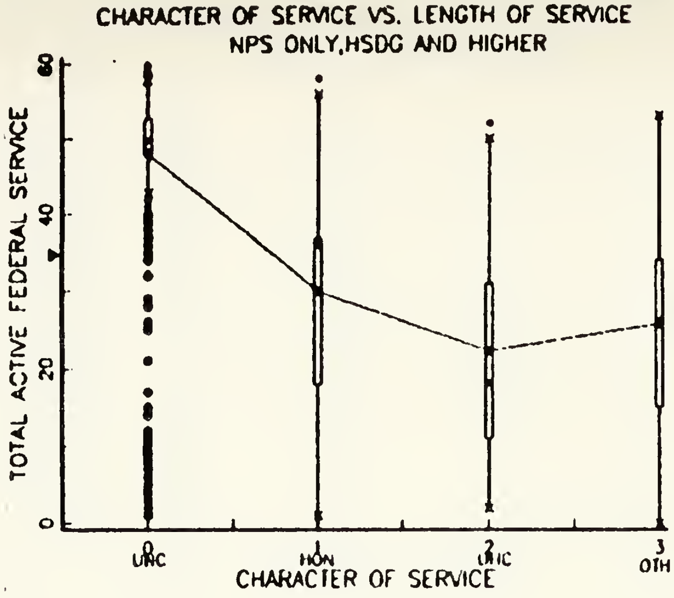
CATEGORY VECTOR :CHCA1 X LABEL
Y :Y←A/8MOD[;1] Y LABEL
SELECTION :(Y≤60)∧(A79MOD[;3]≠4) NO OF ELE

CATEG.	NO.PTS	I•/○-PTS	Y MEAN	Y-DVN	.25	.50	.75
ALL	2932	1	32 33	17 561	17	36	48
0	888	0.30288	48 347	13 874	48	50	54
1	1616	0 55118	27 317	15 862	11	38	36
2	248	0.084584	22 069	13 134	11	19	31
3	180	0.061392	22 322	11.707	13	20	29

CATEGORY VECTOR :CHCA1
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]≠4)∧(B79MOD[;2]≥8)∧

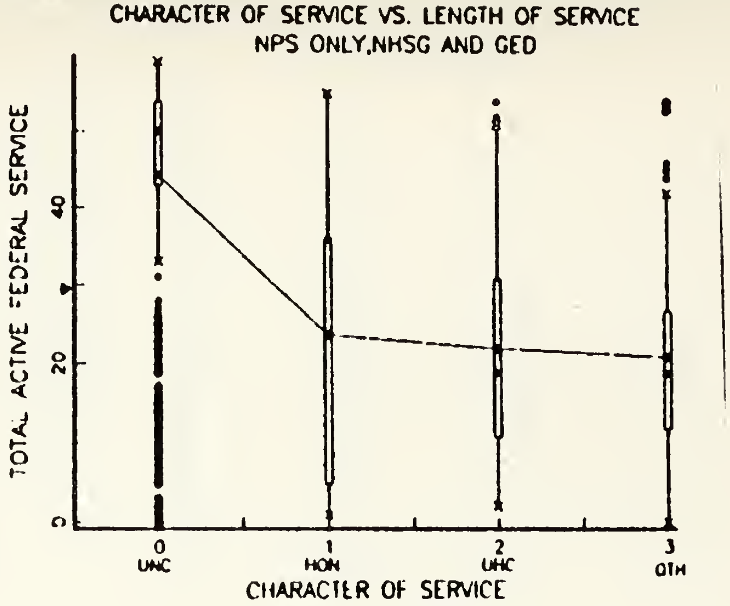
CATEG	NO PTS	I•/○-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1554	1	34 869	18 767	24	36	48
0	471	0.30309	48 212	11.35	48	50	53
1	939	0 60425	29 845	15 513	18	36	37
2	93	0.059846	22 118	13 196	11	18	31
3	51	0 032819	25 569	12 434	15	26	34

Figure H.50 Character of Service vs. Service,III



CATEGORY VECTOR .CHCAT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]≠4)∧(B79MOD[;2]≥6)∧

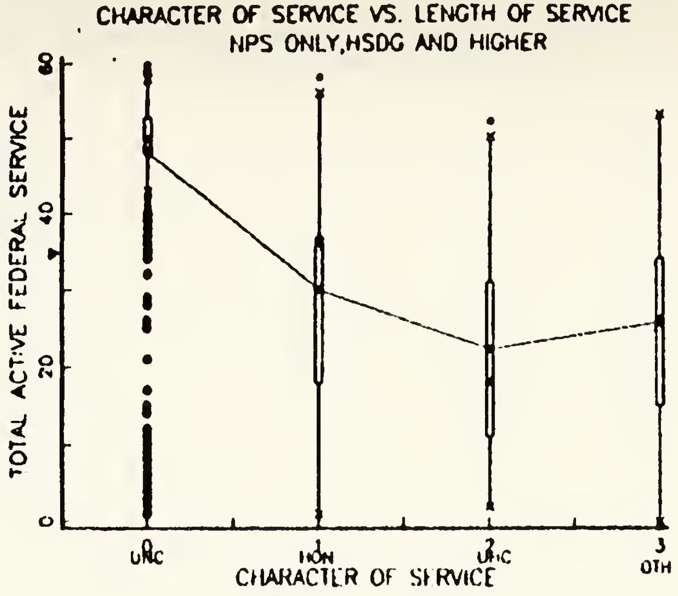
CATEG	NO.PTS	o/o PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1594	1	34 869	16 767	24	36	48
0	471	0 30309	48 212	11.35	48	50	53
1	939	0.60425	29 945	15 513	18	30	37
2	93	0 059846	22.118	13 186	11	18	31
3	51	0 032819	25 569	12.434	15	26	34

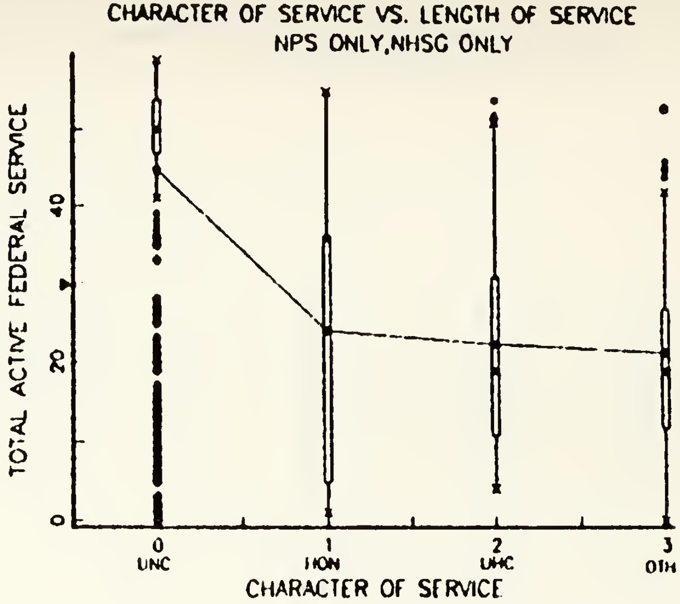


CATEGORY VECTOR .CHCAT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]≠4)∧((B79MOD[;2]≤5)∨

CATEG	NO.PTS	o/o PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1378	1	29 467	17 99	12	36	45
0	417	0 30261	44 24	16	43	50	54
1	677	0 49129	23 674	15 618	5	38	36
2	155	0 11248	22 038	13 097	11	19	31
3	129	0 093614	21 038	11 15	12	18	27

Figure H.51 Character of Service vs. Service,IV





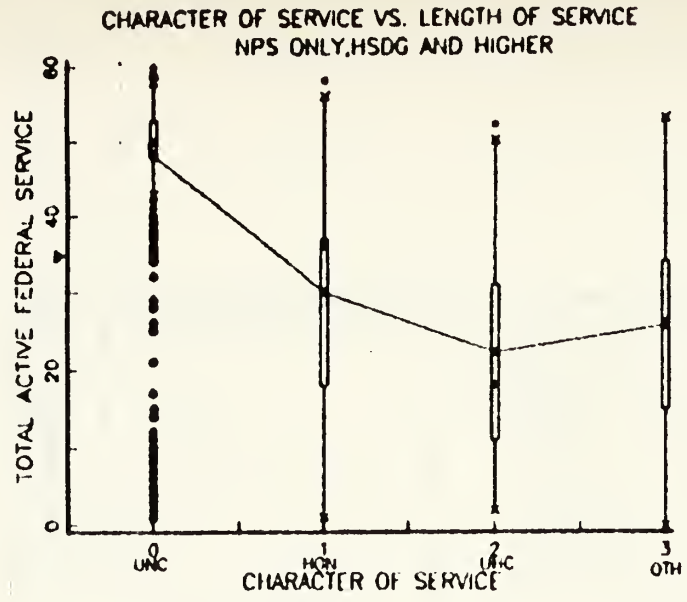
CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]=4)∧(B79MOD[;2]≥6)∧

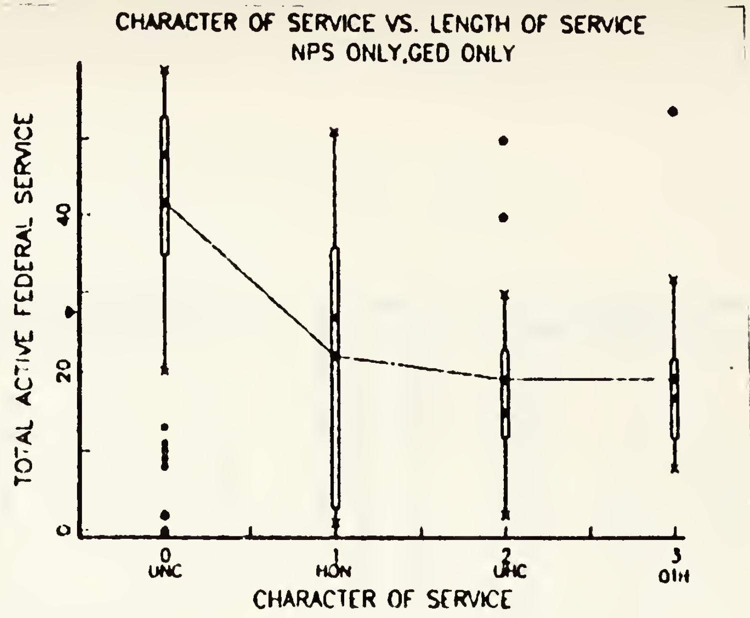
CATEG	NO PTS	+/-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1554	1	34 889	16 767	24	38	48
0	471	0.30309	48 212	11 35	46	50	53
1	939	0 60425	29 945	15 513	18	36	37
2	93	0 059846	22 118	13 198	11	18	31
3	51	0 032818	25.569	12 434	15	26	34

CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]=4)∧(B79MOD[;2]=5)

CATEG	NO PTS	+/-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1197	1	29 773	17 928	12	38	48
0	360	0 30201	44 65	15 805	47	50	54
1	583	0 48909	23 947	15 506	5	36	36
2	141	0 11829	22 326	13 116	11	19	31
3	108	0 090604	21.361	11 282	12	19	27

Figure H.52 Character of Service vs. Service,V





CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]≠4)∧(B79MOD[;2]≥6)∧

CATEG	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	1554	1	34.869	16.767	24	36	48
0	471	0.30309	48.212	11.35	48	50	53
1	939	0.60425	29.945	15.513	18	36	37
2	93	0.059846	22.118	13.196	11	18	31
3	51	0.032819	25.569	12.434	15	26	34

CATEGORY VECTOR :CHCAT
Y :Y←A79MOD[;1]
SELECTION :(Y≤60)∧(A79MOD[;3]≠4)∧(B79MOD[;2]=13)

CATEG	NO PTS	o/o-PTS	Y-MEAN	Y-DVN	.25	.50	.75
ALL	186	1	27.5	18.27	11	31	47
0	57	0.30645	41.848	16.951	35	48	53
1	94	0.50538	21.978	16.192	3	27	38
2	14	0.075269	19.143	12.541	12	15	23
3	21	0.1129	19.381	10.284	12	17	22

Figure H.53 Character of Service vs. Service, VI

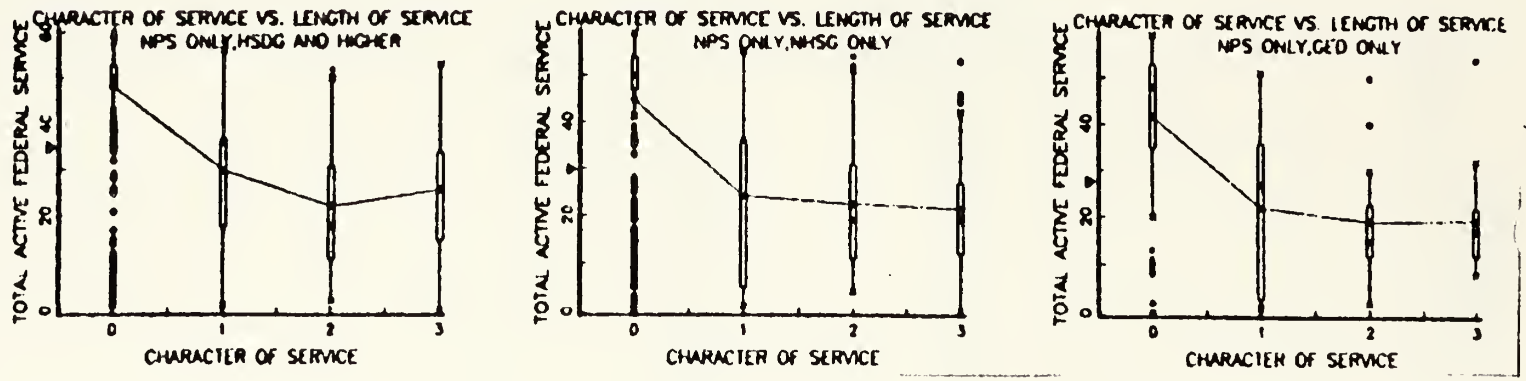


Figure H.54 Character of Service vs. Service, VII

APPENDIX I

APL PROGRAMS TO PERFORM ONE-WAY ANOVA

```
      ∇ANOVA[⎕]∇
     ∇ Z←ANOVA TT;N;K;MAX;BTSS;TSS;TD
[1]    'ANOVA WAS UPDATED 1/3/79, SEE ANOVAHOW FOR CHANGES.'
[2]    →ENTER×⍳1=⍴,TT
[3]    N←+⌿MISS≠T←TT
[4]    MAX←⌈/N
[5]    K←1↓⍴T
[6]    →SUM
[7]   ENTER:INPUT
[8]   SUM:→SKIP×⍳(OPTION=1)∨0=+/+/(T=MISS)
[9]    →0×⍴⍴⎕←'RUN ABORTED; NO MISSING VALUES ALLOWED WHEN OPTION=2.
'
[10]  SKIP:TD←T×(MISS≠T)
[11]   SUMSQ
     ∇
```

```
      ∇SUMSQ[⎕]∇
    ∇ SUMSQ;C;NUMBER;B;BLDF;BLOCK;BLSS;EDF;F;MSBL;MSE;MSR;T;TDF;WTSS;T
       BAR
[1]   T←+⌿TD
[2]   TSS←(+/(+/TD*2))-C←((+/T)*2)÷NUMBER←+/N
[3]   BLSS←(+/(((B←+/TD)*2)÷K))-C
[4]   MSBL←BLSS÷BLDF←MAX-1
[5]   →7×⍳(OPTION=2)
[6]   BLSS←BLDF←0
[7]   WTSS←TSS-BLSS+BTSS←(+/((T*2)÷N))-C
[8]   MSR←BTSS÷K-1
[9]   TDF←NUMBER-1
[10]  F←MSR÷MSE←WTSS÷EDF←TDF-BLDF+K-1
[11]  BLOCK←(B÷K)-TBAR←(+/T)÷NUMBER
[12]  '                          ANOVA TABLE'
[13]  '  SOURCE         DF            SS            MS          F'
[14]  CH←''
[15]  '⎕  TREATMENT⎕,I5,F13.2,F11.2,F8.2' FMT(K-1),BTSS,MSR,F
[16]  →(OPTION=1)/L5
[17]  '⎕  BLOCKS⎕,I8,F13.2,F11.2,F8.2' FMT(MAX-1),BLSS,MSBL,(MSBL÷MSE
)
[18] L5:'⎕  ERROR⎕,I9,F13.2,F11.2' FMT EDF,WTSS,MSE
[19]  '⎕  TOTAL⎕,I9,F13.2' FMT TDF,TSS
[20]  '⎕R-SQUARE = ⎕,F5.3' FMT(BTSS+BLSS)÷TSS
[21]  '⎕OVERALL MEAN = ⎕,F10.2' FMT TBAR
[22]  '⎕TREATMENT EFFECTS ⎕,F6.2' FMT BTSS←(T÷N)-TBAR
[23]  Z←(MISS≠T)×Z←T-(MAX,K)⍴T÷N
[24]  →(OPTION=1)/0
[25]  '⎕BLOCK EFFECTS ⎕,F6.2' FMT BLOCK
[26]  Z←T-TBAR+BLOCK∘.+BTSS
    ∇
```

```
      ∇FMT[⎕]∇
    ∇ OL←E̲ FMT R;S;W;Δ;G;X;T;K;J;M;Q;F;D;N;O;L;B;V;CH;H
[1]   N←Q←1↓M←⍴R←(1⌈¯2↑⍴R)⍴R
[2]   OL←((1=1↑M)↓ 1 0 ×M←M+2↑H←1<⍴CH←C̲H̲,',')⍴Δ←'0123456789.'
[3]   →E×⍳(N←0=N)∨V←1≥⍴S←,E̲
[4]  L0:→~⍳V∨(×F←4×Q=⍴K←⍴X←'  ')∧∨/('A',O←'⎕')∊S
[5]   →(L0+(V←0=⍴S←J↓S)+⍳B=M[2]←1),L̲-(⍳×B←O+.=K),F×~'A'∊K←K,(J←S⍳'
,')↑S
[6]   →E+×⍴S←'TEXT DELIMITER'
[7]   →L3-3××(⍴G←K=K←(K∊¯1↓Δ)/K)⌊W←⍴X←(⍴K←(K⍳O)↑K)↓(-(⌽K)⍳O)↓K
[8]  L̲:→(D←~1↑G←K∊Δ)/L3-2×(⍴K)≠W←1↑O←'XA'∊K←(~K∊' ,')/K
[9]   →L3×⍳(B≠+/G)∧×M[2]←10⊥⌊1-Δ⍳(B←⌊1-G⍳0)↑K
[10]  →L3-⌽O,-(L←'EFI'∊K)/×W←10⊥⌊1-Δ⍳(⌊1-G←B⍳'.')↑B←(1-(⌽G)⍳0)↑K
[11]  A̲←(1↓⍴X←((1⌈⍴A̲)⌊(M[1]-H),W)↑A̲)⌽A̲
[12] L3:→(HD×⍳H∧~'X'∊K),E-⍴X←-W,D←0⍴F←((M-H,0)×1,W)⍴X
[13]  →L4-⍳~1↑L,Q←1↓⍴R←(0 1 ×⍴F←R[;⍳M[2]←Q⌊M[2]⌈Q×V∧D])↓R
[14]  F←F÷10*L←⌊10⍟|F+0=F
[15]  →L3×⍳0=J←+/B←('B'∊K)∧0=F←(⌊0.5+N×,F)÷N←10*D←10⊥⌊1-Δ⍳G↓B
[16] L4:→(⍴1↓⍴L)/F-⍴⍴X←(1 0 ×⍴G←J⍴T\'¯')⍴J←J,O←∨/T←0>F←B/F
[17]  →(×L←(0⌈L×J←'Z'∊K)⌈.×~T←(T+O←1+⌊10⍟1⌈|F)>O+L←W-D+O+~2↓L)/L/
F,F,I
[18]  →E+×⍴S←'FIELD WIDTH'
[19]  →L4+1+⍳((J[2]←L∨.<0)+O←1+10⌈.≤|L←(B/,L)+T←10=|F)>W-D+O+3
[20]  T←~J←F[T/⍳1↑J]←L←⍴1⍴X←'E','+0¯'[J⍴2-×L],Δ[1+⍉(O⍴10)⊤|L]
[21] F:→(J∨2≥D×~'T'∊K)/I,N←⍴X←Δ[11,1+⍉(D⍴10)⊤⌈N×1||F],X
[22]  D←,(-N)↑(D⌈.×⍉X[;2+D]≠1↑Δ)∘.<D←⍳D-1
[23]  X←N⍴X,X[D/⍳⍴X←,X]←' '
[24] I:→(J←J∨0=+/O←0⌈L-O)/I+⍴D←⍴F←G,Δ[1+⍉(L⍴10)⊤L|F]
[25]  F←D⍴(,O←G⌽O)\(,O←D∘.<(-G)⌽⍳L+G←1↓⍴G)/,F
[26]  →HD-⍳J∨L←~'L'∊K,F[T/⍳D←1↑X←⍴F←F,X;]←'*'
[27]  F←X⍴(,⌽O)\(,O←~X↑~O)/,F
[28]  →(~H)/E-N←1,D←0⍴F←B⍀(D,X←W×1-2×L)↑F
[29] HD:CH←(⍴K←(¯1↑D←0,(M[2]⌊⍴D)⍴D←(','=CH)/⍳⍴CH)⍴CH)⌽CH
[30]  D←,(M[2],X)↑ 0 ¯1 ↓(M[2],B)⍴(,⌽D∘.≥⍳B←⌈/D←1↓D-¯1⌽D)\K
[31]  →(L0-V∧×Q),⍴OL←OL,((1=1↑M)↓M×1,W)⍴D,,F
[32] E:K←'NO VALID E, I, OR F PHRASE'
[33]  (⍞'FMT PROBLEM     ',K),⍞(1,⍴S)⍴S
    ∇
```

APPENDIX J

INPUT SCREEN FOR IBM GRAFSTAT CDF PROGRAM

```
                         CUMULATIVE DISTRIBUTIONS

TYPE OF PLOT              : SURVIVOR
DATA                      : X
DISTRIBUTION              : (TO BE FIT)  PARAMETER SPECIFICATION: Δ
CENSORING                 : 0            ESTIMATION METHOD      : Δ
                                         CONFIDENCE LEVEL(S)    : .95

SELECTION                 : Δ
FREQUENCIES               . 1
K-S BOUNDS                : N            K-S CONFIDENCE LEVEL(S):  95
NORMALIZE (Y/N)           : Y

PLOT HEADER (IN QUOTES)   : Δ
SCREEN HEADER (IN QUOTES): ''
X AXIS LABEL (IN QUOTES) : Δ
Y AXIS LABEL (IN QUOTES) : Δ
POSITION                  : 1
SCALE X-AXIS              : LIN          SCALE Y-AXIS  : LIN 0 YMAX
PARTIAL PLOT              : 1 1 1        LINE TYPES    : 1 1
AXES AND GRID CONTROL     : 0 1 0 0

ENTER=GO         PF:1=HELP  2=VIEW GRAPHICS(3279)  3=RETURN   4=WRITE ON SCREEN
CLEAR=DEFAULT       5=LAST RESPONSES     6=ERASE      7=AXIS CONTROL     8=TABLE
     RESPONSES      9=OUTPUT  10=STORE/RETRIEVE  11=INTO APL  12=SCREEN DISPLAY
```

APPENDIX K

SURVIVOR CURVES FOR 3 YEAR ENLISTEES

This Appendix contains survivor curves for FY79 3 year-obligated enlistees for the six candidate explanatory variables listed in Table XVII below. Tabular summaries of the analysis and discussion of the analysis is provided in Chapter 4 of this thesis.

TABLE XVII

Candidate Explanatory Variables

Sex
Race
Mental Category
Marital Status/Number of Dependents
Age
Military Occupational Skill

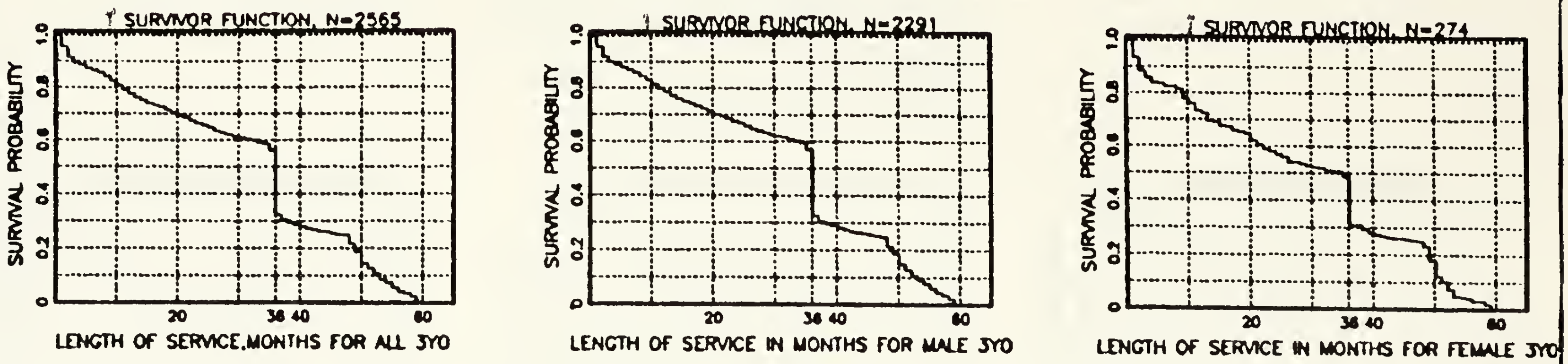


Figure K.1 Sex

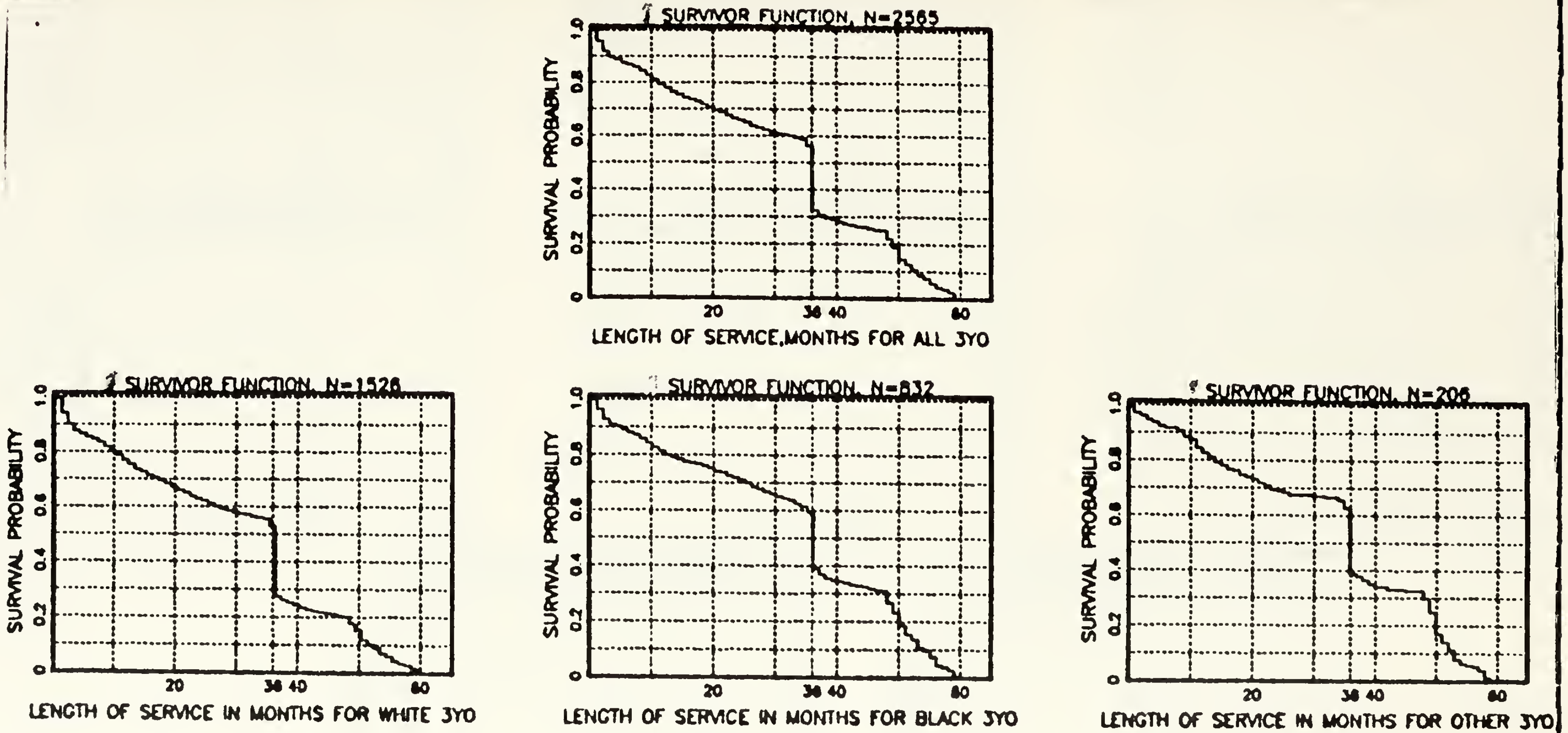


Figure K.2 Race

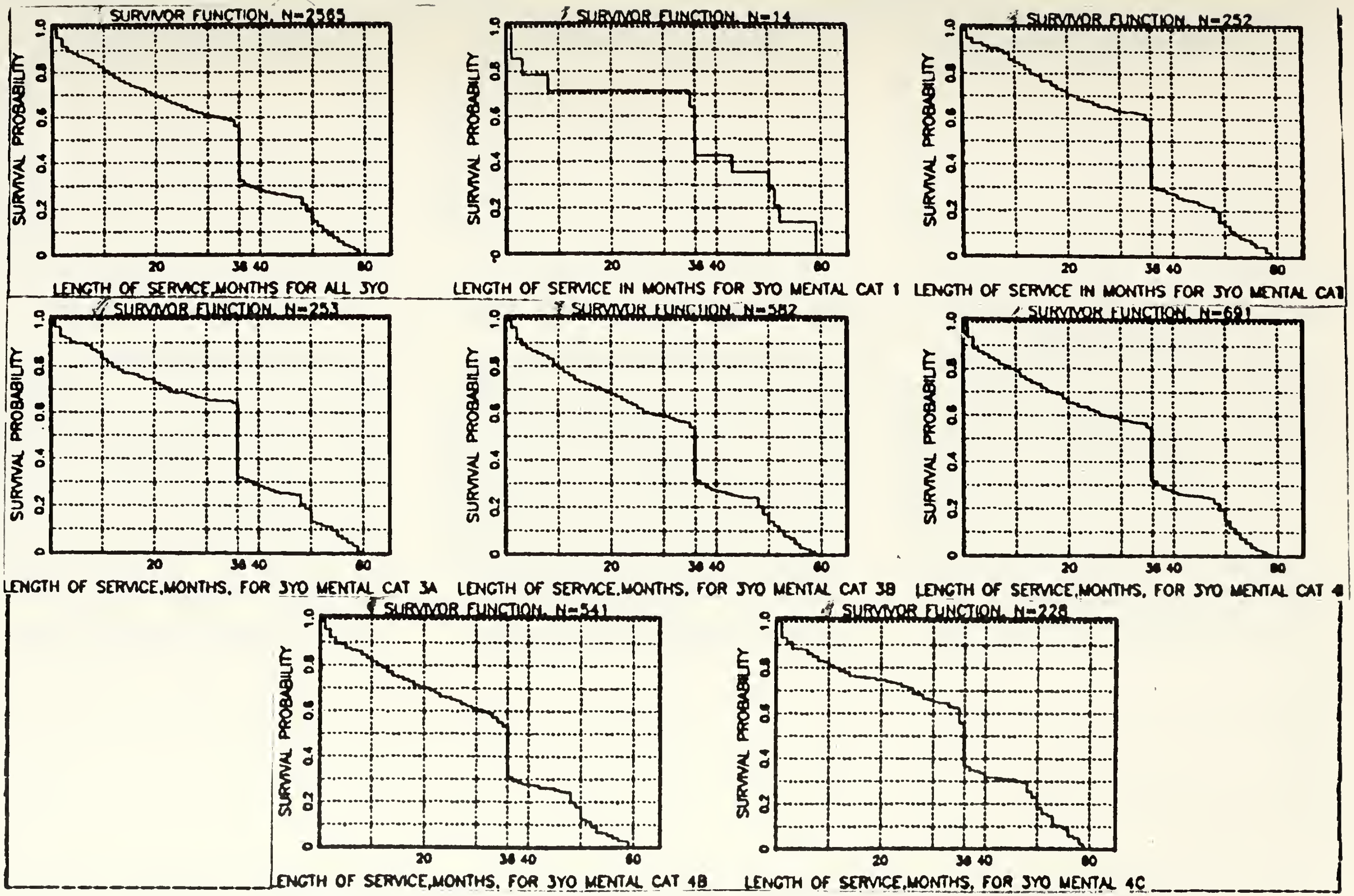


Figure K.3 Mental Category

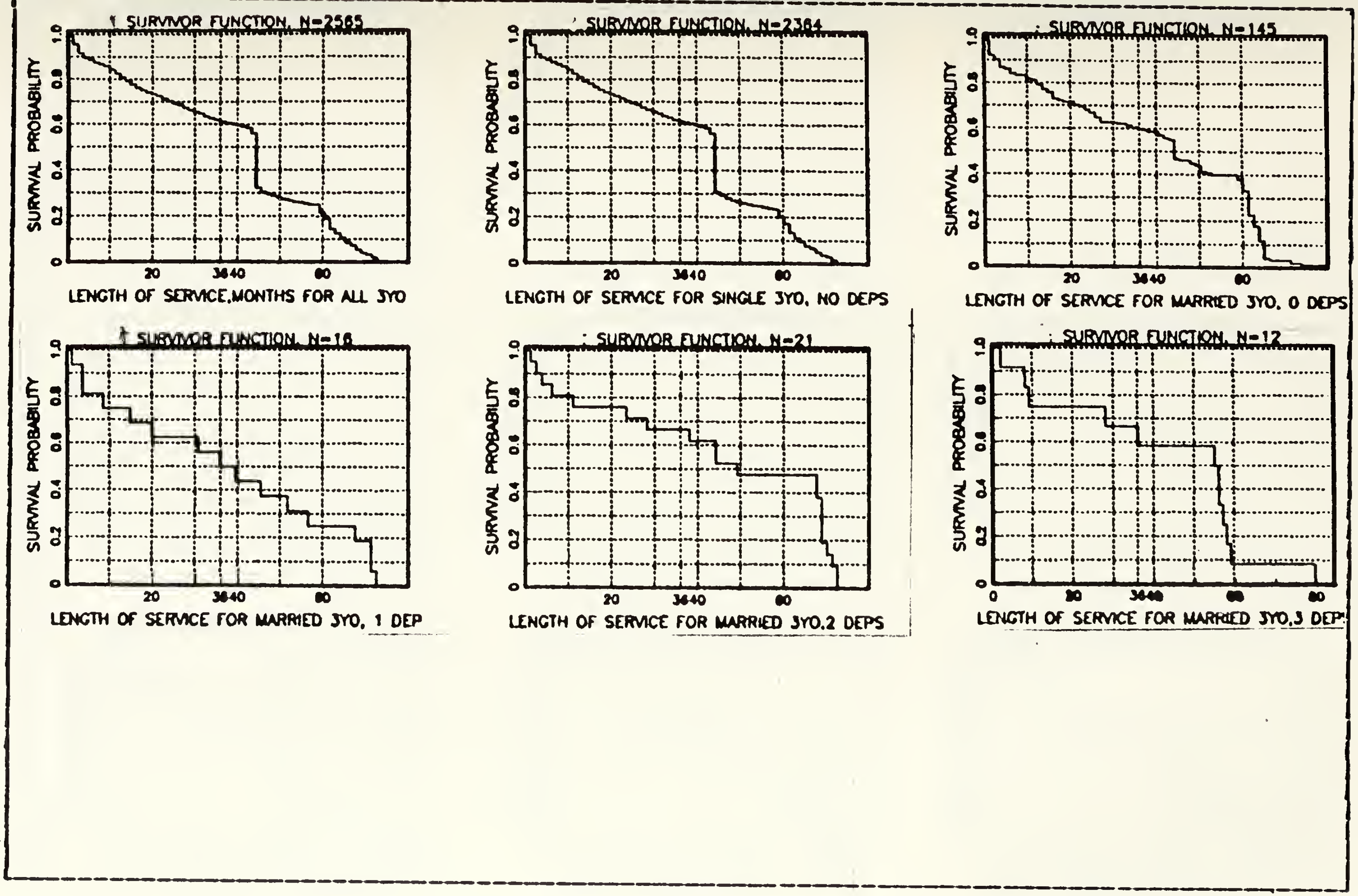


Figure K.4 Marital Status/ Number of Dependents

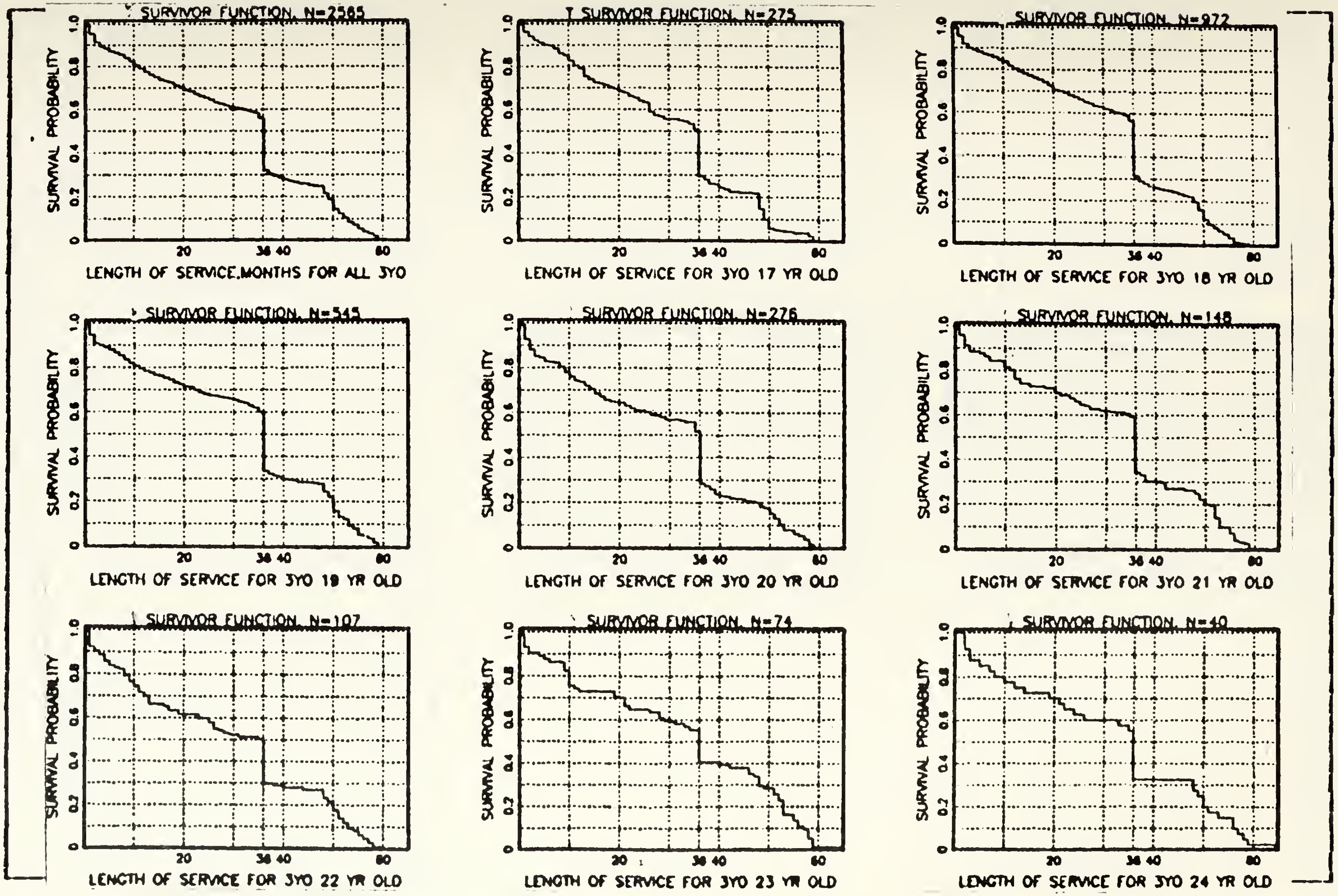


Figure K.5 Age I

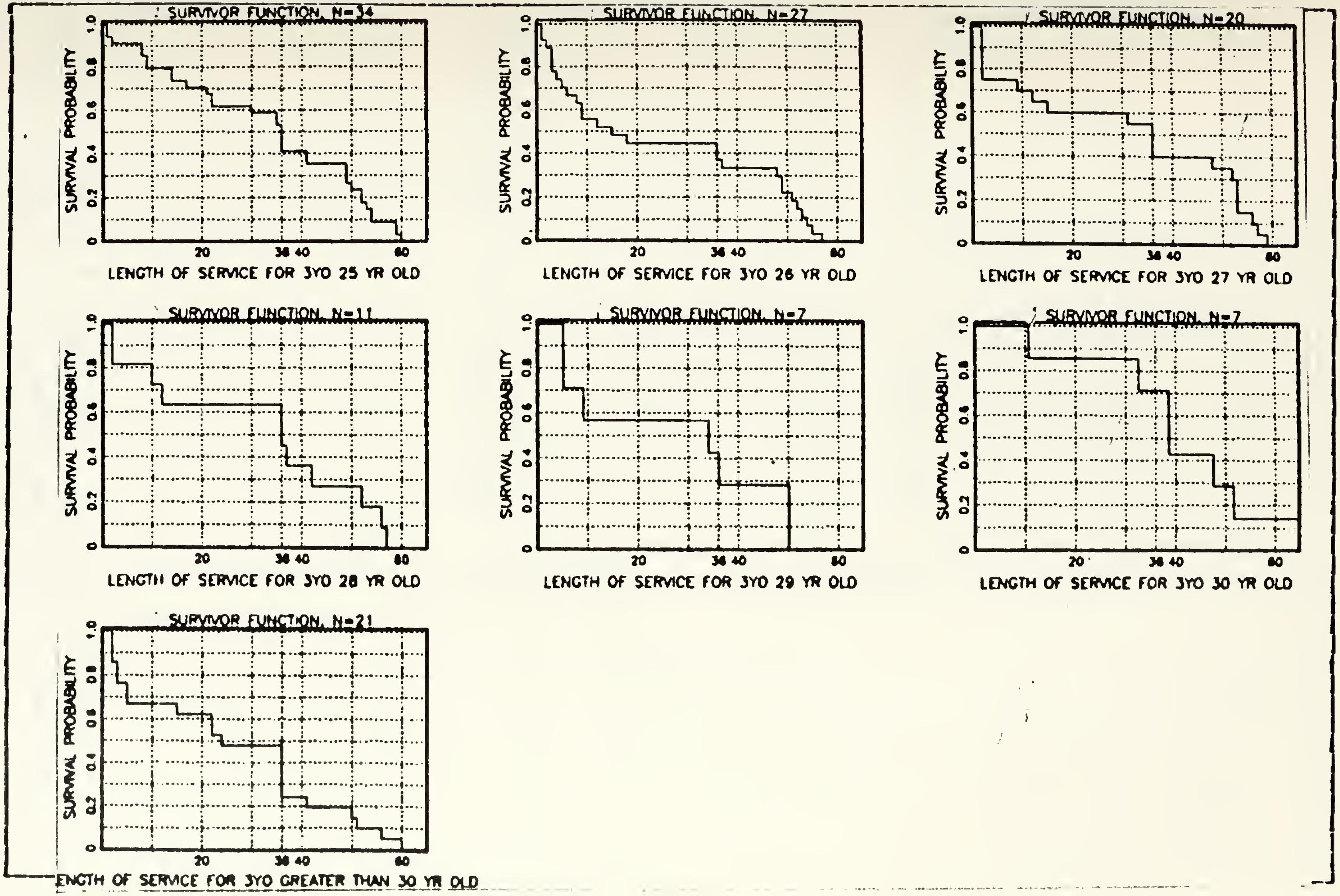


Figure K.6 Age II

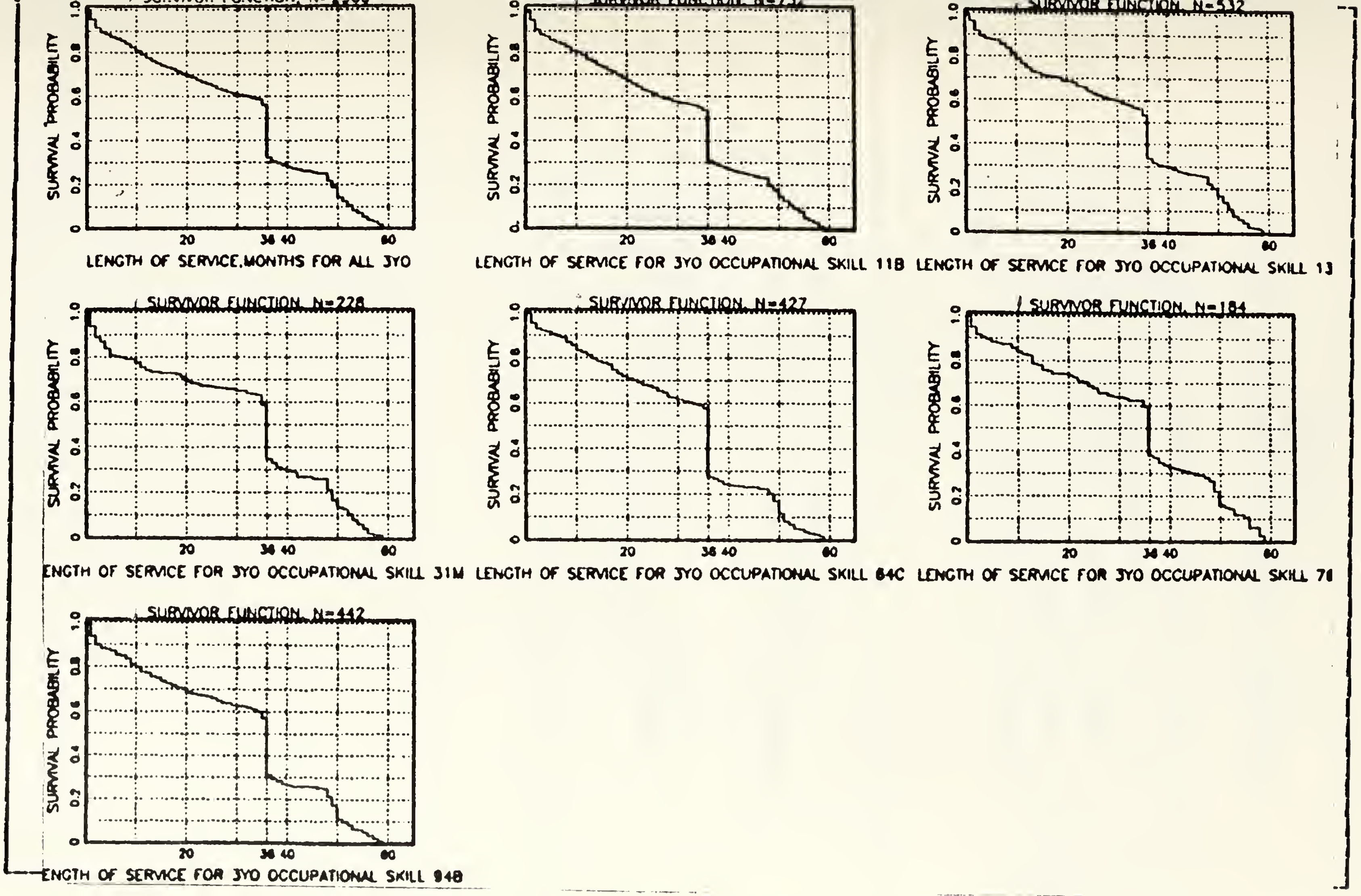


Figure K.7 Military Occupational Skill

APPENDIX L

SURVIVOR CURVES FOR 4 YEAR ENLISTEES

This Appendix contains survivor curves for FY79 4 year-obligated enlistees for the six candidate explanatory variables listed in Table XVII below. Tabular summaries of the analysis and discussion of the analysis is provided in Chapter 4 of this thesis.

TABLE XVIII

Candidate Explanatory Variables

Sex
Race
Mental Category
Marital Status/Number of Dependents
Age
Military Occupational Skill

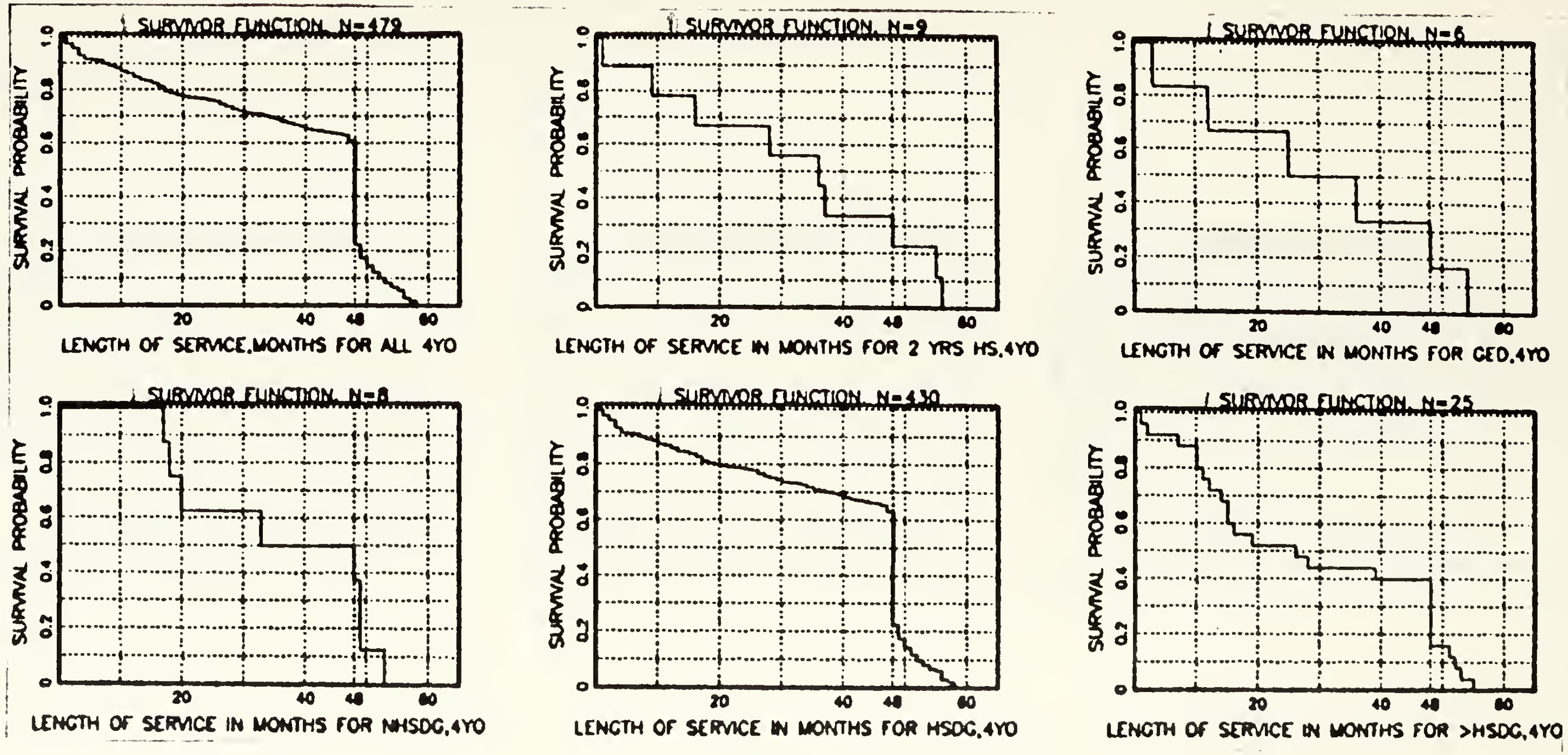


Figure L.1 Education Level

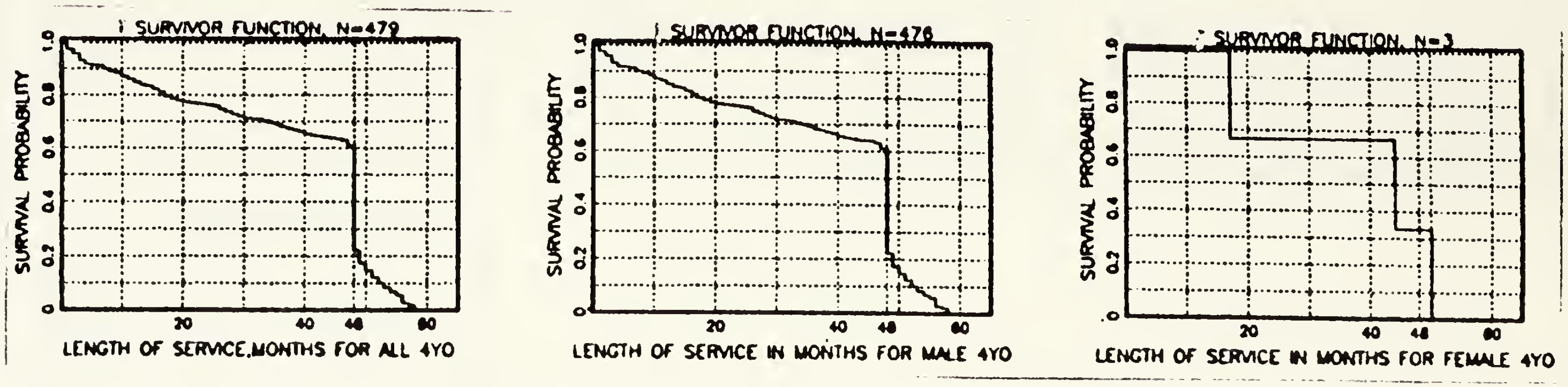


Figure L.2 Sex

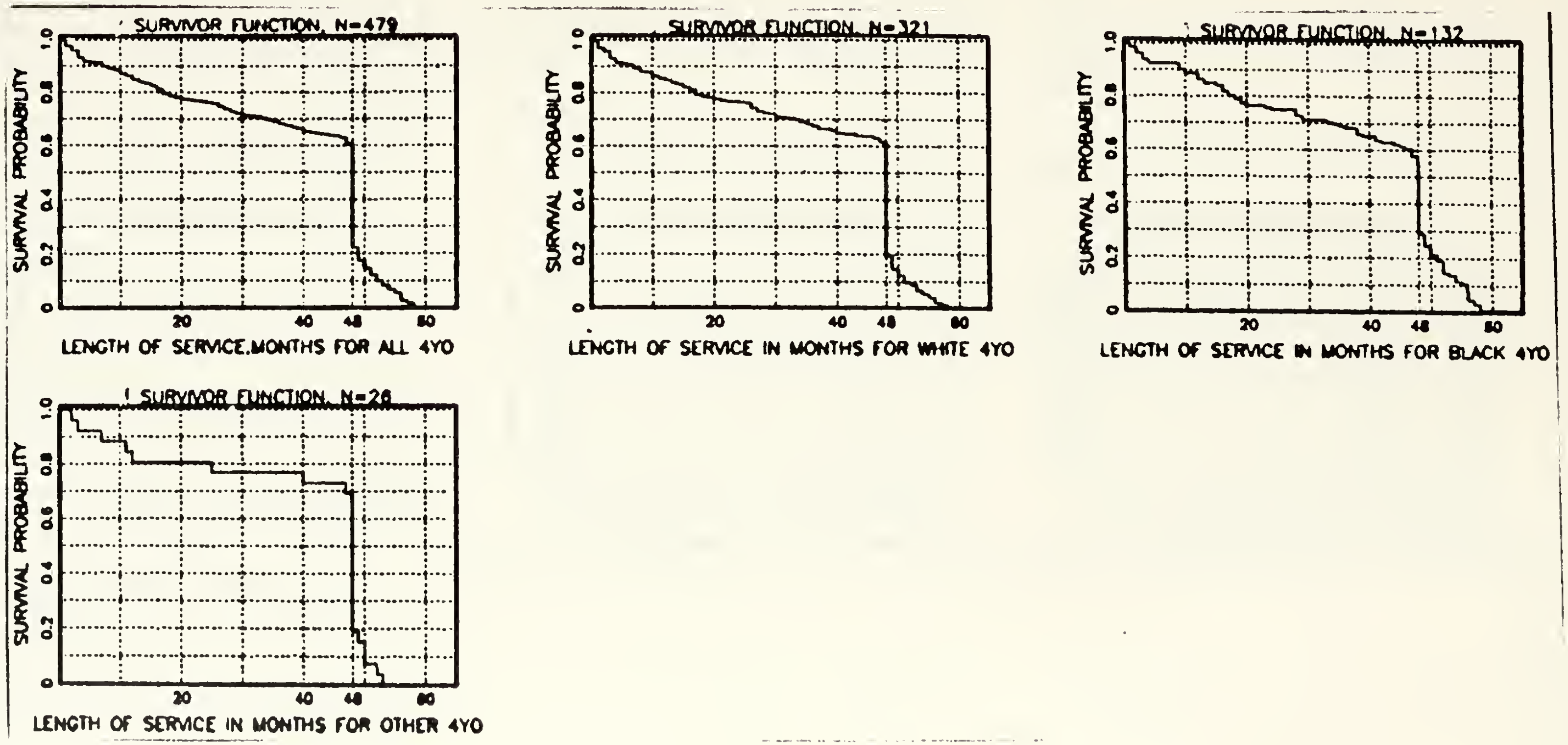


Figure L.3 Race

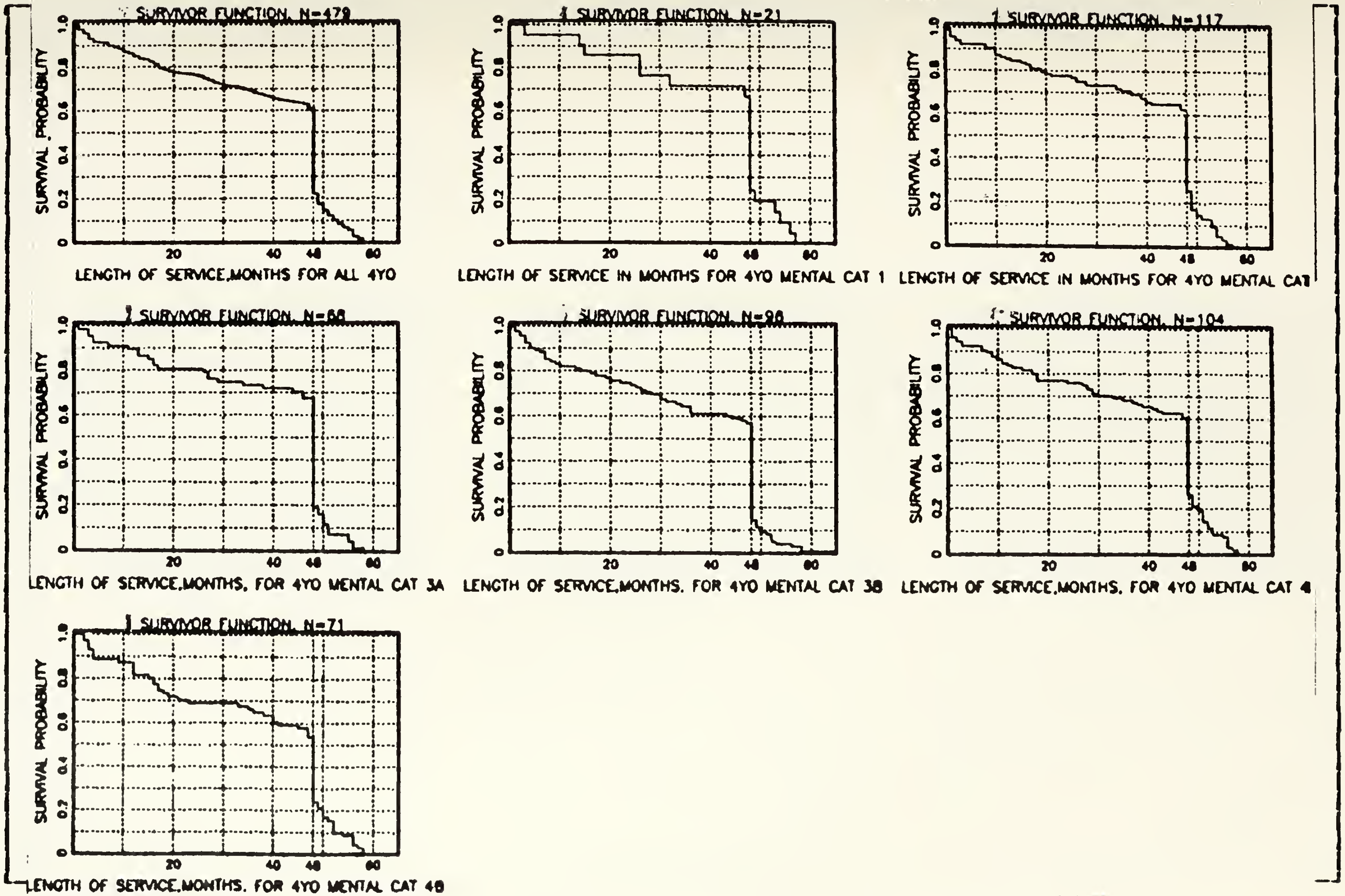


Figure L.4 Mental Category

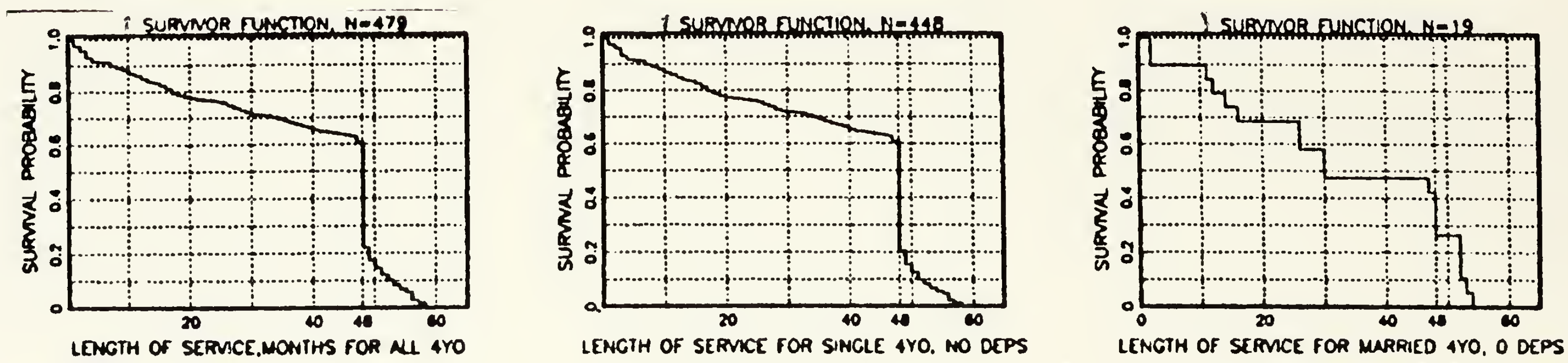


Figure L.5 Marital Status/ Number of Dependents

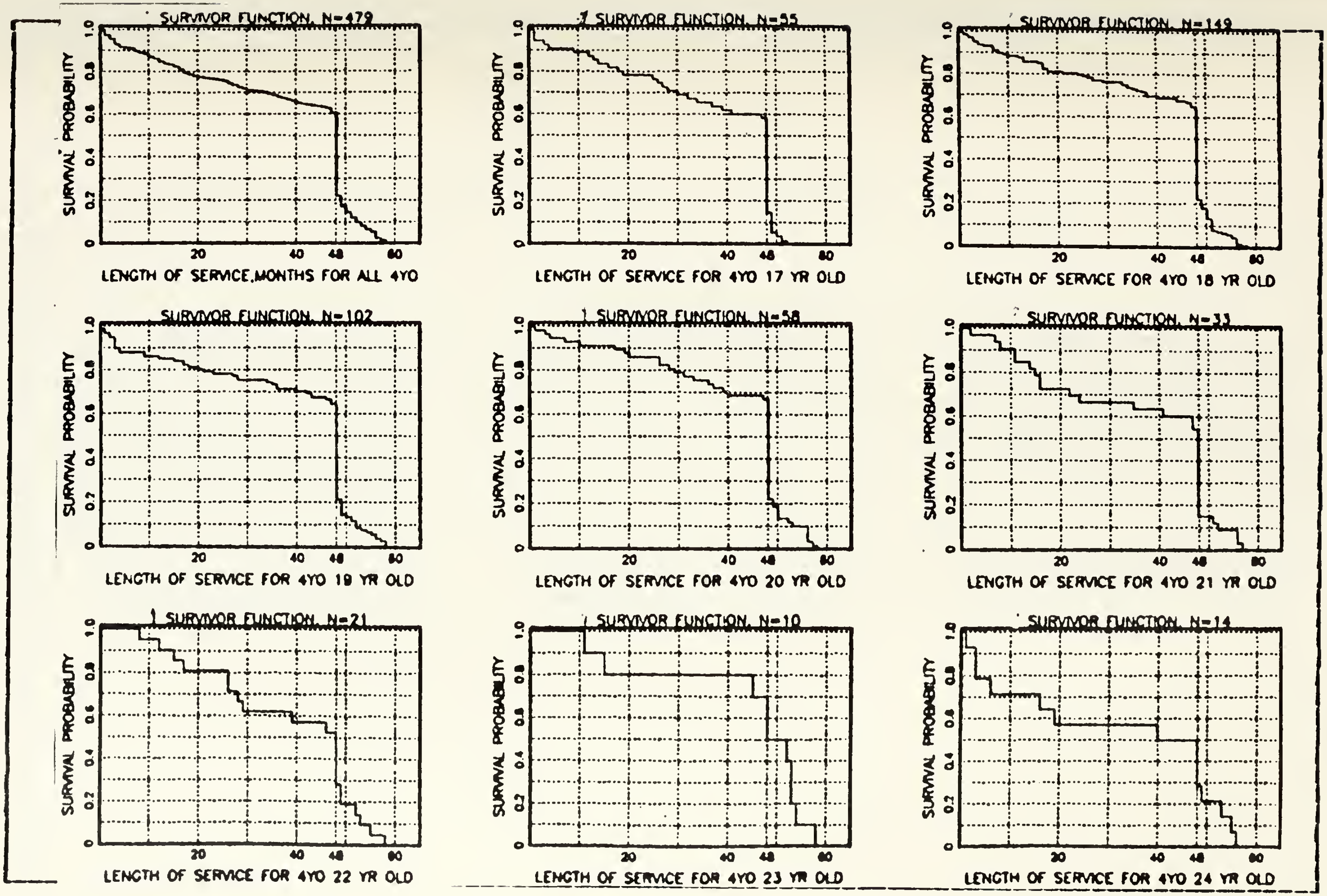


Figure L.6 Age I

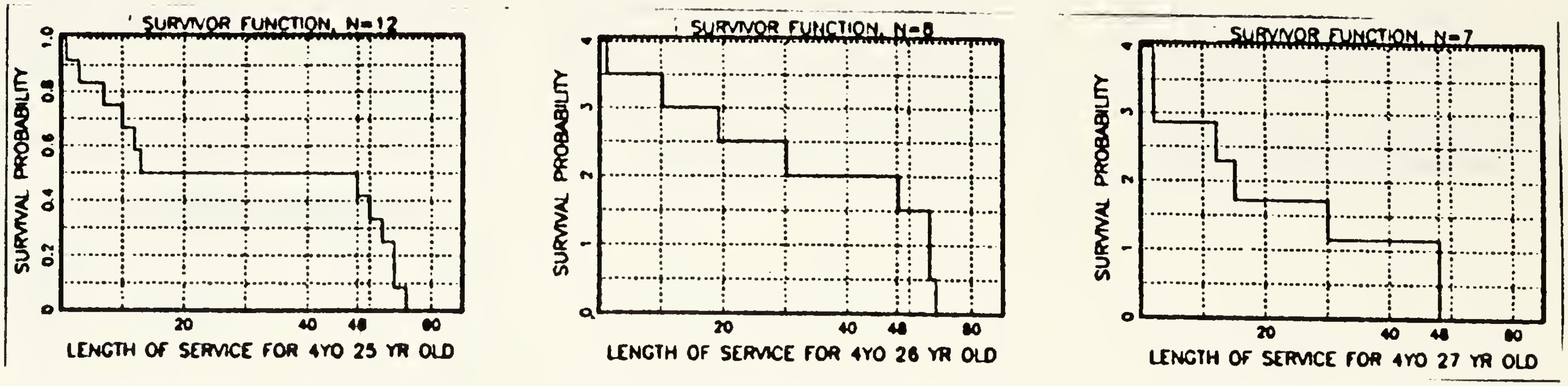


Figure L.7 Age II

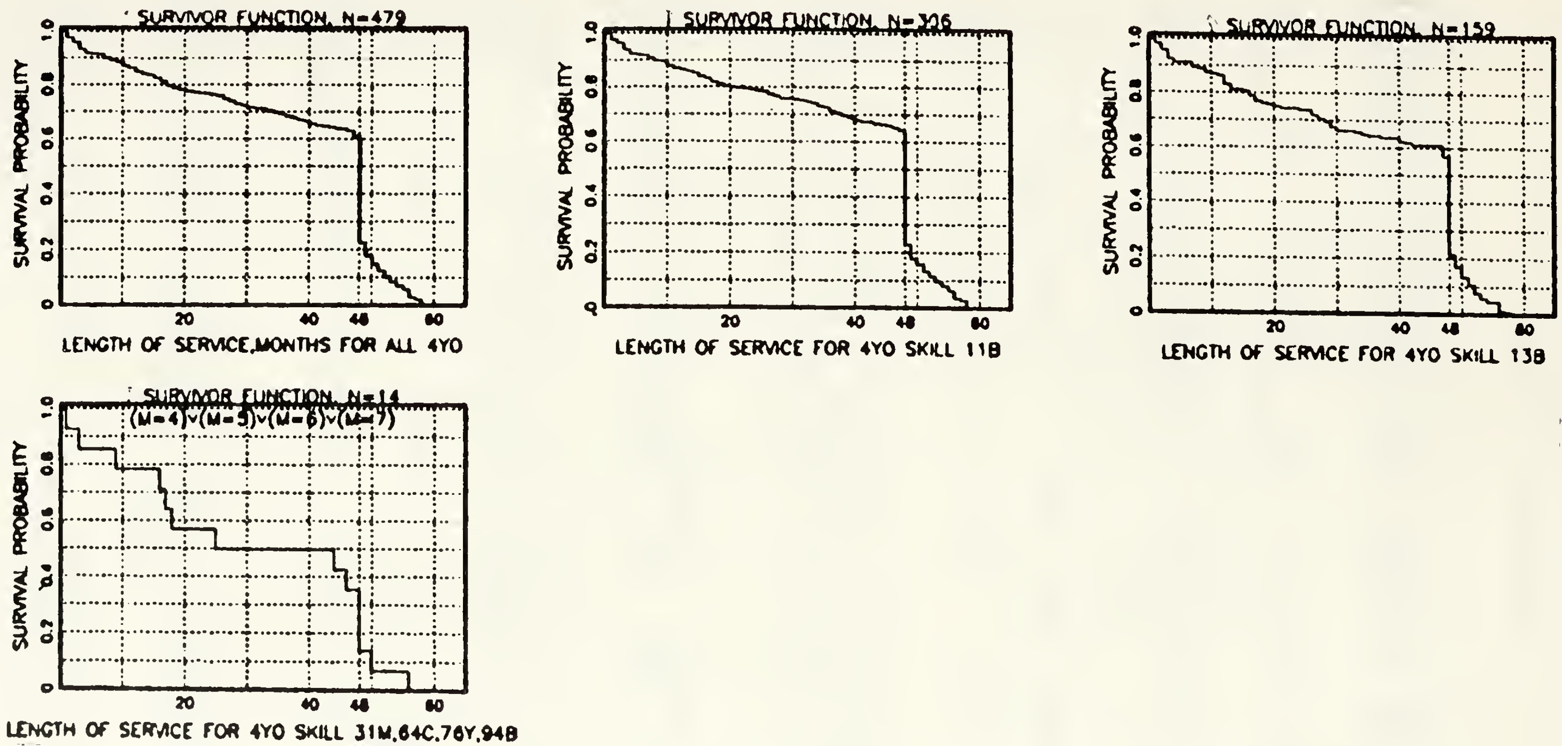


Figure L.8 Military Occupational Skill

LIST OF REFERENCES

1. Rand Corporation Report R-2986-MRAL, The Military Enlistment Process, What Happens and Can It Be Improved?, by S.E. Berryman, R.M. Bell, and W.Lisowski, May 1983.

2. "The All Volunteer Army: A Decade of Success," All Volunteer, v. 36, December 1983.

3. Rand Corporation Report R-1450 ARPA, Military Manpower and the All Volunteer Force, by Richard V.L. Cooper, September 1977.

4. Comptroller of the Army, The Army Budget, Fiscal Year 1985, January 1984.

5. Telephone conversation with CPT Wilton Ham, analyst, USAREC, 28 Jan 1984

6. Foch, C. and Nikki King, "The Gatekeepers: First-Term Enlisted Attrition Policies and Practices, Paper presented at Conference on First Term Enlisted Attrition sponsored by OSD and ONR, Leesburg, Virginia, 4-7 April 1977.

7. Center for Naval Analysis Professional Paper 177, Predicting Attrition: A Test of Alternative Approaches, by R.F. Lockman and J.T. Warner, March 1977.

8. Bovard, James. "Slow the Military Retirement Gravy Train", The Wall Street Journal, v.CX, no. 78, 20 April 1984.

9. Pamphlet prepared by USAREC, USAREC 1983 Report

10. Center for Naval Analyses, Report CRC 345, Prediction of Attrition From Military Service, by John T. Warner, September 1978.

11. Defense Manpower Data Center, Internal Working Papers, "DMDC File Format Sheet" and "Coding and Data Element Description", February 1982. These documents are available from DMDC upon request.

12. Chambers, John M., et. al. Graphical Methods for Data Analysis, Belmont, California: Wadsworth International Group, 1983.

13. Johnson, Malcom. Draftsman Displays, A Graphical Technique for Exploratory Data Analysis, MSOR Thesis, Naval Postgraduate School, Monterey, California, 1984.

14. Marshall, K.T., and Grinold, R.C. Manpower Planning Models, North Holland, 1977.

15. Larson, Harold J. Introduction to Probability Theory and Statistical Inference, 3rd edition, New York: John Wiley and Sons, 1982.

16. Conover, W.J. Practical Non-Parametric Statistics, 2d Edition, New York: Wiley and Sons, 1980.

17. Dixon, W.J. BMDP Statistical Software, Berkeley: University of California Press, 1981.

18. Olkin, Ingram, et.al. Probability Models and Applications, New York: Macmillan Publishing Company. Inc., 1980.

19. Kalbfleisch, John D., and Ross L. Prentice. The Statistical Analysis of Failure Time Data, New York: Wiley and Sons, 1980.

20. McCullagh P., and J.A. Nelder. Generalized Linear Models, New York: Chapman and Hall, 1983.

21. SUGI Supplemental Users Guide, Cary, North Carolina: SAS Institute, 1983.

22. Tukey, John W. Exploratory Data Analysis, Reading, Massachusetts: Addison-Wesley Publishing Co., 1977

23. Gaver, Donald P., Statistical Methods, Some Old, Some New: A Tutorial Survey, NPS paper prepared for ONR, Naval Postgraduate School, January 1981.

24. Computer Center Newsletter, Naval Postgraduate School, May 18, 1983.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA. 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, CA. 93943	2
3.	Dr. P.A.W. Lewis, Code 55Lw Department of Operations Research Naval Postgraduate School Monterey, CA 93943	6
4.	Dr. George W. Thomas, Code54Te Naval Postgraduate School Monterey, CA 93943	6
5.	Dr. Philip Heidelberger IBM Thomas J. Watson Research Center Yorktown Heights, NY 10598	1
6.	Dr. P.D. Welch IBM Thomas J. Watson Research Center Yorktown Heights, NY 10598	1
7.	Headquarters U.S. Army Recruiting Command ATTN: Program Analysis and Evaluation Directorate Fort Sheridan, Illinois 60037	1
8.	Deputy Under Secretary of the Army (Operations Research) Room 2E621, The Pentagon Washington, DC 20310	1
9.	CPT David A. Thomas 1807 Catawba Street Fayetteville, N.C. 28303	5